REPRESENTATIONS OF A COMPLEX POINT BY PAIRS OF ORDERED REAL POINTS *

BY

W. C. GRAUSTEIN

1. **Introduction.** A complex point z in a complex space of n dimensions, since it depends on $2n$ real parameters, can be represented by a pair of real points, X and Y, of the space. It is reasonable to restrict the choice of X and Y by the following conditions:

A. The points X and Y representing z shall be ordered: $X \to Y$.

B. If z is a real point, X and Y shall coincide in z.

C. If z is imaginary, X and Y shall be distinct; if $X \to Y$ represents z, $Y \to X$ shall represent $\bar{z}$, the point conjugate-imaginary to z.

D. The coördinates of X and Y, referred to a rectangular coördinate system, shall be analytic functions, without singularities in the finite domain, of the coördinates of z and $\bar{z}$, referred to this system.

E. The representation of z by $X \to Y$ shall be invariant under a chosen group of real point transformations; that is, the ordered points $X' \to Y'$ into which the ordered points $X \to Y$ are carried by an arbitrary transformation of the group shall represent the point z' into which z is carried by the transformation.

If $X \to Y$ is a representation of z which satisfies these conditions, $Y \to X$ is also a representation of z satisfying them. The two representations shall be called *inverse*.

There are two pairs of inverse representations satisfying the prescribed conditions which are well known: those of Laguerre in the plane, whereby z is represented by the real points of intersection, taken in the one or the other order, of the minimal lines through z and $\bar{z}$; and those of Marie in the plane and in space, in which the real points of the representation lie on the line $z\bar{z}$, have the same midpoint as z and $\bar{z}$, and are at a distance apart whose square is the negative of the square of the distance $z\bar{z}$.† The two Marie representations are invariant under the group of affine transformations, whereas those of Laguerre are unchanged by the direct circular transformations of the plane.

This paper deals with the determination of all the representations $X \to Y$

* Presented to the Society, December 31, 1919.

† For a review of the literature concerning these representations, see E. Study, *Vorlesungen über ausgewählte Gegenstände der Geometrie, erstes Heft; Ebene analytische Kurven und zu ihnen gehörige Abbildungen*, Leipzig, 1911; J. L. Coolidge, *Geometry in the Complex Domain* (in press).

of z which satisfy conditions A to D, and are invariant, in compliance with E, under a given group. In § 2 the problem is formulated analytically for an arbitrary group. In the subsequent sections, 3–7, it is solved for various subgroups of the projective group, for the projective group itself, and for the group of direct circular transformations of the plane. Only in the case of the projective group are no representations obtained.

Given a group, G, of direct transformations and the set, S, of corresponding indirect transformations, it is of importance to determine, among the representations invariant under G, those which are invariant also under S, and secondly, those which are inverted, i.e., carried into their inverses, by S. It turns out that all the representations invariant under G are invariant also under S, except when G is the group of direct circular transformations, or when G is either the group of motions or of direct transformations of similarity in the plane.* In the latter cases certain representations of the totality invariant under G are unchanged by S, whereas certain others are inverted by S (§ 8). Moreover, it is only in these same cases that the representations are of exceptional nature in comparison with those obtained for the same group applied in a space of different dimensionality.

2. **Analytic formulation of the problem.** Let $(z_1, z_2, \cdots, z_n)$, where

$$z_i = x_i + iy_i \qquad (i = 1, 2, \cdots, n),$$

be the coördinates, referred to a rectangular cartesian system of axes, of the point z, and let the coördinates of X and Y, referred to the same system, be respectively $(X_1, X_2, \cdots, X_n)$ and $(Y_1, Y_2, \cdots, Y_n)$.

Since the latter sets of coördinates are, by condition D, real analytic functions of $z_1, z_2, \cdots, z_n, \bar{z}_1, \bar{z}_2, \cdots, \bar{z}_n$, they are also real analytic functions of the real variables $x_1, y_1, x_2, y_2, \cdots, x_n, y_n$. It is convenient, in light of B, to write these functions as follows:

$$X_i = x_i + F_i(x_1, y_1, \cdots, x_n, y_n),$$
$$Y_i = x_i + G_i(x_1, y_1, \cdots, x_n, y_n) \qquad (i = 1, 2, \cdots, n).$$

By C,

$$G_i(x_1, y_1, \cdots, x_n, y_n) = F_i(x_1, -y_1, \cdots, x_n, -y_n) \qquad (i = 1, 2, \cdots, n).$$

Conditions A to D are thus covered by demanding that X_i and Y_i be of the form

$$\begin{aligned} X_i &= x_i + F_i(x_1, y_1, \cdots, x_n, y_n), \\ Y_i &= x_i + F_i(x_1, -y_1, \cdots, x_n, -y_n) \qquad (i = 1, 2, \cdots, n), \end{aligned} \tag{1}$$

where $F_1, F_2, \cdots, F_n$ are real functions of the real variables $x_1, y_1, \cdots, x_n, y_n$, which are analytic for all finite sets of values, vanish identically when

$$y_1 = y_2 = \cdots = y_n = 0,$$

* In this summary of exceptions, one relatively unimportant case has been omitted.

and are furthermore such that the system of equations

$$F_i(x_1, y_1, \cdots, x_n, y_n) = F_i(x_1, -y_1, \cdots, x_n, -y_n)$$

has no real solutions in $y_1, y_2, \cdots, y_n$ except the solution $0, 0, \cdots, 0$.

Let the equations

$$z_i' = T_i(z_1, z_2, \cdots, z_n) \qquad (i = 1, 2, \cdots, n) \tag{2}$$

represent an arbitrary group of real point transformations. In order that the representation $X \to Y$ of the point z be invariant under this group, it is necessary and sufficient that

$$\begin{aligned} &(a)\ x_i' + F_i(x_1', y_1', \cdots, x_n', y_n') \\ &\quad = T_i(x_1 + F_1(x_1, y_1, \cdots, x_n, y_n), \cdots, \\ &\qquad x_n + F_n(x_1, y_1, \cdots, x_n, y_n)), \\ &(b)\ x_i' + F_i(x_1', -y_1', \cdots, x_n', -y_n') \\ &\quad = T_i(x_1 + F_1(x_1, -y_1, \cdots, x_n, -y_n), \cdots, \\ &\qquad x_n + F_n(x_1, -y_1, \cdots, x_n, -y_n)) \quad (i = 1, 2, \cdots, n). \end{aligned} \tag{3}$$

Thus our problem, finally formulated, consists in solving the system of functional equations (3) for the functions $F_1, F_2, \cdots, F_n$, subject to the conditions stated under (1).

If the transformations (2) are linear, as will frequently be the case,

$$x_i' = T_i(x_1, x_2, \cdots, x_n), \qquad y_i' = T_i^0(y_1, y_2, \cdots, y_n) \qquad (i = 1, 2, \cdots, n),$$

where T_i^0 is T_i minus its absolute term. Changing the signs of $y_1, y_2, \cdots, y_n$ then changes the signs of $y_1', y_2', \cdots, y_n'$. Consequently, the two systems of equations in (3) are, in this case, each equivalent to the single system

$$\begin{aligned} F_i(x_1', y_1', \cdots, x_n', y_n') = T_i^0(&F_1(x_1, y_1, \cdots, x_n, y_n), \\ &\cdots, F_n(x_1, y_1, \cdots, x_n, y_n)) \quad (i = 1, 2, \cdots, n). \end{aligned} \tag{4}$$

3. **Representations invariant under the group of translations.** Inasmuch as the group of real translations is a subgroup of each of the other groups under which the invariance of representations is to be considered, it is advantageous to deal with it first. Equations (4) then become

$$F_i(x_1 + a_1, y_1, \cdots, x_n + a_n, y_n) = F_i(x_1, y_1, \cdots, x_n, y_n) \qquad (i = 1, 2, \cdots, n).$$

Since $a_1, a_2, \cdots, a_n$, the parameters of the translation, are arbitrary, the system is satisfied if and only if $F_1, F_2, \cdots, F_n$ are functions of $y_1, y_2, \cdots, y_n$ alone.

The representations $X \to Y$ of z invariant under the group of real translations

are given by

$$\text{I} \qquad X_i = x_i + F_i(y_1, y_2, \cdots, y_n), \quad Y_i = x_i + F_i(-y_1, -y_2, \cdots, -y_n) \qquad (i = 1, 2, \cdots, n),$$

where $F_1, F_2, \cdots, F_n$ *are arbitrary real analytic functions which vanish for* $y_1 = y_2 = \cdots = y_n = 0$ *and are such that the system of equations*

$$F_i(y_1, y_2, \cdots, y_n) = F_i(-y_1, -y_2, \cdots, -y_n) \quad (i = 1, 2, \cdots, n)$$

has no real solutions other than $0, 0, \cdots, 0$.

In the case of a group of linear transformations containing the group of translations, (4) can now be replaced by the system of equations

$$(5) \qquad F_i(y_1', y_2', \cdots, y_n') = T_i^0(F_1(y_1, y_2, \cdots, y_n), \cdots, F_n(y_1, y_2, \cdots, y_n)) \quad (i = 1, 2, \cdots, n),$$

in which $x_1, x_2, \cdots, x_n, x_1', x_2', \cdots, x_n'$ no longer appear.*

4. **The group of homothetic transformations.** The representation $X \to Y$ is invariant under the group of real homothetic transformations

$$z_i' = \rho z_i + a_i, \qquad \rho \neq 0 \qquad (i = 1, 2, \cdots, n),$$

if and only if the system (5), formed for this group, namely

$$F_i(\rho y_1, \rho y_2, \cdots, \rho y_n) = \rho F_i(y_1, y_2, \cdots, y_n) \quad (i = 1, 2, \cdots, n),$$

is satisfied. But then, regardless of whether ρ is unrestricted or takes on only positive values, $F_1, F_2, \cdots, F_n$ must be linear and homogeneous in $y_1, y_2, \cdots, y_n$.

The representations $X \to Y$ *of* z *invariant under the group of either direct or general homothetic transformations are*

$$\text{II} \qquad X_i = x_i + \sum_{j=1}^{n} k_{ij} y_j, \qquad Y_i = x_i - \sum_{j=1}^{n} k_{ij} y_j \quad (i = 1, 2, \cdots, n),$$

where the k_{ij} *are real constants subject only to the restriction that the* n*-rowed determinant* $|k_{ij}|$ *does not vanish.*

* The equations (5) are also necessary and sufficient conditions that the real analytic point transformation

$$(a) \qquad y_i' = F_i(y_1, y_2, \cdots, y_n) \qquad (i = 1, 2, \cdots, n)$$

be commutative with each transformation of the group of real linear transformations

$$(b) \qquad y_i' = T_i^1(y_1, y_2, \cdots, y_n) \qquad (i = 1, 2, \cdots, n)$$

leaving the origin invariant. Accordingly, as we solve (5) in the subsequent paragraphs for various groups (*b*), we determine at the same time, for each particular group (*b*), the group of transformations (*a*), each of which is commutative with every transformation (*b*). Thus we find, for example, in § 6, that the group of transformations commutative with the group of motions about the origin in the plane is

$$y_1' = y_1 R_1(r^2) + y_2 R_2(r^2), \qquad y_2' = -y_1 R_2(r^2) + y_2 R_1(r^2),$$

where $r^2 = y_1^2 + y_2^2$ and R_1, R_2 are analytic functions of r^2, not both zero.

The condition $|k_{ij}| \neq 0$ is necessary and sufficient, not merely that X and Y coincide only when z is real, but also that the $2n$ equations II can be solved uniquely for $x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_n$ and hence the coördinates of z found in terms of those of X and Y. In other words, the correspondence between the ordered point pairs $X \to Y$ and the points z under a representation II is one-to-one without exception in the finite domain.

5. **The affine and projective groups.** In order that the representation II be invariant under the n transformations

$$z_q' = -z_q, \qquad z_i' = z_i \qquad (q = 1, 2, \cdots, n),$$

where in each transformation i takes on all the integral values from 1 to n except the fixed value q, it is necessary that $k_{ij} = 0$, $i \neq j$. If the resulting representation is to be unchanged by each of the transformations

$$z_q' = z_r, \qquad z_r' = z_q, \qquad z_i' = z_i \qquad (q \neq r),$$

where in each transformation q and r are fixed and i takes on the remaining values from 1 to n, it follows that $k_{11} = k_{22} = \cdots = k_{nn}$. A necessary condition, therefore, that II be invariant under the affine group is that $k_{ij} = 0$, $i \neq j$, and $k_{ii} = k$ for all i. That this condition is also sufficient can be easily verified.

If, instead of all real affine transformations, merely those for which the determinant of the coefficients is positive had been admitted, the result would have been the same.

The representations $X \to Y$ of z invariant under the group of either direct or general affine transformations are

III $$X_i = x_i + ky_i, \qquad Y_i = x_i - ky_i \qquad (i = 1, 2, \cdots, n),$$

where k is an arbitrary constant, $\neq 0$.

The two inverse representations for which $k = \pm 1$ are *the representations of Marie.*

If III is to be invariant under the projective group, it must be unchanged by the transformation

$$z_1' = \frac{1}{z_1}, \qquad z_i' = \frac{z_i}{z_1} \qquad (i = 2, \cdots, n).$$

But then the first of the equations (3) necessitates that $k^2 = -1$.

There is no representation $X \to Y$ of z which is invariant under the projective group.

6. **The group of motions.** The treatment in this case varies with the dimensionality of the space.

A. $n = 1$. The group of motions is the group of translations. According to I, *the invariant representations are*

IVA $$X_1 = x_1 + y_1 \Phi(y_1), \qquad Y_1 = x_1 - y_1 \Phi(y_1),$$

where $\Phi(y_1)$ is a real analytic function and the equation $\Phi(y_1) = -\Phi(-y_1)$ is satisfied by no real value of y_1 except perhaps $y_1 = 0$.

B. $n = 2$. The representation must be of the form

$$X_i = x_i + F_i(y_1, y_2), \qquad Y_i = x_i + F_i(-y_1, -y_2) \quad (i = 1, 2), \tag{6}$$

and be invariant under the transformations

$$z_1' = \cos\phi\, z_1 - \sin\phi\, z_2 + a_1, \qquad z_2' = \sin\phi\, z_1 + \cos\phi\, z_2 + a_2.$$

That is, $F_1(y_1, y_2)$, $F_2(y_1, y_2)$ must be solutions of the system of equations, obtained from (5),

$$\begin{aligned} F_1(y_1', y_2') &= \cos\phi\, F_1(y_1, y_2) - \sin\phi\, F_2(y_1, y_2), \\ F_2(y_1', y_2') &= \sin\phi\, F_1(y_1, y_2) + \cos\phi\, F_2(y_1, y_2), \end{aligned} \tag{7}$$

where

$$y_1' = \cos\phi\, y_1 - \sin\phi\, y_2, \qquad y_2' = \sin\phi\, y_1 + \cos\phi\, y_2.$$

For $\phi = \pi$, equations (7) become the conditions that F_1, F_2 be odd functions of y_1, y_2. Consequently, if F_i is written in the form

$$F_i(y_1, y_2) = \frac{F_i(y_1, y_2) + F_i(y_1, -y_2)}{2} + \frac{F_i(y_1, y_2) + F_i(-y_1, y_2)}{2} \qquad (i = 1, 2),$$

the first quotient on the right is even in y_2, odd in y_1 and hence identically zero for $y_1 = 0$, whereas the second is even in y_1, odd in y_2 and so identically zero for $y_2 = 0$. Accordingly,

$$F_i(y_1, y_2) = y_1 F_i'(y_1, y_2) + y_2 F_i''(y_1, y_2) \qquad (i = 1, 2), \tag{8}$$

where each of the functions F_1', F_1'', F_2', F_2'' is analytic and is even in y_1, in y_2, and hence in y_1, y_2.

In the equations obtained from (7) by substituting for F_1 and F_2 their values as given by (8), replace ϕ by $-\phi$, y_1 by $-y_1$, and combine by addition and subtraction each of the resulting equations with that from which it arose. The system obtained is

$$\begin{aligned} y_1' F_1'(y_1', y_2') &= y_1 \cos\phi\, F_1'(y_1, y_2) - y_2 \sin\phi\, F_2''(y_1, y_2), \\ y_2' F_2''(y_1', y_2') &= y_1 \sin\phi\, F_1'(y_1, y_2) + y_2 \cos\phi\, F_2''(y_1, y_2), \\ y_1' F_2'(y_1', y_2') &= y_1 \cos\phi\, F_2'(y_1, y_2) + y_2 \sin\phi\, F_1''(y_1, y_2), \\ y_2' F_1''(y_1', y_2') &= -y_1 \sin\phi\, F_2'(y_1, y_2) + y_2 \cos\phi\, F_1''(y_1, y_2). \end{aligned} \tag{9}$$

These equations reduce, for $\phi = \pi/2$, to

$$F_2'(y_1, y_2) = -F_1''(y_2, y_1), \qquad F_2''(y_1, y_2) = F_1'(y_2, y_1). \tag{10}$$

When the values for F_2' and F_2'' given by (10) are substituted in (9) and y_2 is

set equal to zero, the four resulting equations yield immediately the identities

$$F_1'(y_1 \cos\phi, y_1 \sin\phi) = F_1'(y_1 \sin\phi, y_1 \cos\phi),$$
$$F_1''(y_1 \cos\phi, y_1 \sin\phi) = F_1''(y_1 \sin\phi, y_1 \cos\phi),$$

in the independent variables y_1 and ϕ. Consequently, $F_1'(y_1, y_2)$ and $F_1''(y_1, y_2)$ are symmetric functions of y_1, y_2, and (10) becomes

$$(11) \qquad F_2'(y_1, y_2) = -F_1''(y_1, y_2), \qquad F_2''(y_1, y_2) = F_1'(y_1, y_2).$$

But the system (9) then reduces to

$$F_1'(y_1', y_2') = F_1'(y_1, y_2), \qquad F_1''(y_1', y_2') = F_1''(y_1, y_2).$$

That is, F_1' and F_1'' are absolute invariants of the group of rotations about the origin, and hence must be of the form

$$(12) \qquad F_1' = R_1(r^2), \qquad F_1'' = R_2(r^2),$$

where

$$r^2 = y_1^2 + y_2^2.$$

The necessary conditions (8), (11), (12) that $F_1(y_1, y_2)$, $F_2(y_1, y_2)$ be solutions of (7) are shown sufficient by direct substitution.

The representations $X \to Y$ of z invariant under the group of motions in the plane are

$$\text{IVB} \qquad \begin{aligned} X_1 &= x_1 + y_1 R_1(r^2) + y_2 R_2(r^2), & Y_1 &= x_1 - y_1 R_1(r^2) - y_2 R_2(r^2), \\ X_2 &= x_2 - y_1 R_2(r^2) + y_2 R_1(r^2), & Y_2 &= x_2 + y_1 R_2(r^2) - y_2 R_1(r^2), \end{aligned}$$

where $r^2 = y_1^2 + y_2^2$, and R_1, R_2 are real functions of r^2, analytic for all finite values and not vanishing simultaneously except perhaps when $r^2 = 0$.

C. $n \geqq 3$. In case $n = 3$, the functions F_1, F_2, F_3 in the representation

$$(13) \qquad X_i = x_i + F_i(y_1, y_2, y_3), \qquad Y_i = x_i + F_i(-y_1, -y_2, -y_3) \qquad (i = 1, 2, 3)$$

are to be determined from the system of equations

$$(14) \qquad F_i(y_1', y_2', y_3') = \sum_{j=1}^{3} a_{ij} F_j(y_1, y_2, y_3) \qquad (i = 1, 2, 3),$$

where

$$y_i' = \sum_{j=1}^{3} a_{ij} y_j \qquad (i = 1, 2, 3),$$

and the determinant $|a_{ij}|$ is orthogonal and has the value $+1$.

In case of an arbitrary rotation about the axis of z_3, $y_3' = y_3$ and y_1, y_2 are transformed as for $n = 2$. The last equation of (14) reduces to

$$F_3(y_1', y_2', y_3) = F_3(y_1, y_2, y_3),$$

and the first two to the equations (7), with the variable y_3 added in each

function. Consequently,

$$
\begin{aligned}
F_1 &= y_1 F_3'(y_1^2+y_2^2, y_3) + y_2 F_3''(y_1^2+y_2^2, y_3),\\
F_2 &= -y_1 F_3''(y_1^2+y_2^2, y_3) + y_2 F_3'(y_1^2+y_2^2, y_3),\\
F_3 &= F_3'''(y_1^2+y_2^2, y_3).
\end{aligned}
\tag{15}
$$

Corresponding respectively to the rotations about the axes of z_1 and z_2, there are two other sets of equations like (15) and obtainable from (15) by permuting the subscripts 1, 2, 3 cyclicly. From the three sets of equations it follows that $F_1'' = F_2'' = F_3'' = 0$; for example, since $F_1 = F_1'''(y_2^2+y_3^2, y_1)$, F_1 is an even function of y_2, whence, by the first equation of (15), $F_3'' = 0$. Therefore

$$F_i = y_i F_j'(y_i^2+y_k^2, y_j) = y_i F_k'(y_i^2+y_j^2, y_k),$$

where i, j, k are to be permuted cyclicly through the values 1, 2, 3. Thus

$$F_1' = F_2' = F_3' = R(r^2), \quad \text{where} \quad r^2 = y_1^2+y_2^2+y_3^2,$$

and (13) becomes

$$X_i = x_i + y_i R(r^2), \qquad Y_i = x_i - y_i R(r^2) \qquad (i = 1,2,3).$$

Substitution shows that the functions F_i in this representation satisfy (14).

This result for $n = 3$ can be generalized by mathematical induction to hold for the case $n > 3$.

The representations $X \to Y$ of z invariant under the group of motions in n-dimensional space, where $n \geqq 3$, are

IVC $\qquad X_i = x_i + y_i R(r^2), \qquad Y_i = x_i - y_i R(r^2) \quad (i = 1, 2, \cdots, n),$

where

$$r^2 = y_1^2 + y_2^2 + \cdots + y_n^2,$$

and $R(r^2)$ is a real function analytic for all finite values and not zero except perhaps when $r^2 = 0$.

For the group of motions and reflections the result incorporated in IVC *holds for all values of n.* To establish the truth of this statement for $n \geqq 3$, it is sufficient to show that the representation IVC is invariant under a reflection in one of the coördinate hyperplanes. If $n = 2$, the representation must be of the form IVB and be invariant under $z_1' = z_1, z_2' = -z_2$; then $R_2(r^2) = 0$, and conversely. Finally, IVA is invariant under $z_1' = -z_1$ if and only if $\Phi(y_1)$ is an even function of y_1, i.e., is of the form $R(y_1^2)$.

7. **The group of direct transformations of similarity.** Necessary and sufficient that the representation $X \to Y$ be unchanged by the direct transformations of similarity in the cases $n = 1$, $n = 2$, $n \geqq 3$ is the requirement that IVA, IVB, IVC be respectively invariant under an arbitrary stretching from the origin of positive ratio. According to § 4, the arbitrary functions involved in the representations IV are, then, all constants, and conversely.

The representations $X \to Y$ of z invariant under the group of direct transformations of similarity are

VA
$$X_1 = x_1 + ky_1, \qquad Y_1 = x_1 - ky_1, \qquad k \neq 0, \quad \text{if} \quad n = 1,$$

VB
$$X_1 = x_1 + ky_1 + ly_2, \qquad Y_1 = x_1 - ky_1 - ly_2,$$
$$X_2 = x_2 - ly_1 + ky_2, \qquad Y_2 = x_2 + ly_1 - ky_2, \qquad k^2 + l^2 \neq 0, \quad \text{if} \quad n = 2,$$

VC
$$X_i = x_i + ky_i, \qquad Y_i = x_i - ky_i, \qquad k \neq 0, \quad i = 1, 2, \cdots, n, \quad \text{if} \quad n \geqq 3.$$

Except in the plane, these representations are identical with the representations III invariant under the affine group.

For the group of direct and indirect transformations of similarity the representations are for all n given by VC *and are identical with those peculiar to the affine group.*

The two inverse representations VB, for which $k = 0$ and $l = \pm 1$, namely

VI
$$X_1 = x_1 \pm y_2, \qquad Y_1 = x_1 \mp y_2,$$
$$X_2 = x_2 \mp y_1, \qquad Y_2 = x_2 \pm y_1,$$

are *the representations of Laguerre.* From their geometrical definition (§ 1), it is evident that these representations are invariant under the direct circular transformations. It can be shown that they are the only representations with this property.

8. **Representations inverted by a set of transformations.** Given a group, G, of direct transformations of a certain type and the set, S, of indirect transformations of the same type, to determine, among the representations invariant under G, those which, though not invariant under S, are merely inverted by the transformations of S.

It is geometrically evident that a Laguerre representation is inverted by an indirect circular transformation. Of the groups G other than the circular group, every one under which invariant representations exist is linear and contains the group of translations. Hence the representation must be of the form I. In order that it be inverted by a linear transformation (2), it is necessary and sufficient that the functions F_i satisfy the system of equations

$$F_i(-y_1', -y_2', \cdots, -y_n') = T_i^0(F_1(y_1, y_2, \cdots, y_n), \cdots, F_n(y_1, y_2, \cdots, y_n)) \qquad (i = 1, 2, \cdots, n), \tag{16}$$

In §§ 4–7 it has been shown that, when G is any one of the following groups of direct transformations: homothetic group, affine group, group of motions

($n \geqq 3$), similarity group ($n \neq 2$), all the representations invariant under G are also invariant under S. It is true also that no one of the representations IVA, invariant under the group of motions for $n = 1$, is inverted by any transformation of S, as can readily be proved by applying (16) with $z_1' = -z_1$ as the transformation used; but in this case, not all the representations are invariant under S (§ 6).

If G is the group of motions (or of direct transformations of similarity) in the plane, the functions F_i to be used in (16) are as in IVB (or VB). It is sufficient to choose as the transformation a reflection in a coördinate axis. Then equations (16) reduce to $R_1 = 0$ (or $k = 0$).

Of the representations IVB *invariant under the group of motions in the plane, only those for which* $R_2 = 0$ *are invariant under reflections, whereas only those for which* $R_1 = 0$ *are inverted by every reflection.*

Of the representations VB *invariant under the direct transformations of similarity, only those for which* $l = 0$ *are invariant under the indirect transformations of similarity, whereas only those for which* $k = 0$ *are inverted by every such transformation.*

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

NON-LOXODROMIC SUBSTITUTIONS AND GROUPS IN N DIMENSIONS*

BY

EDWARD B. VAN VLECK

Linear substitutions

$$x' = \frac{ax + b}{cx + d} \tag{1}$$

are commonly divided into two classes, loxodromic and non-loxodromic, and the non-loxodromic substitutions are further classified as elliptic, parabolic, and hyperbolic. So far as I am aware, no corresponding classification of projective substitutions in $n - 1$ nonhomogeneous or n homogeneous variables has been suggested.† The classification which is introduced below is based upon the formulation of a projective substitution under the shape of a homogeneous substitution with determinant $+1$. If the substitution (1) is written in homogeneous form

$$x_1' = ax_1 + bx_2, \qquad x_2' = cx_1 + dx_2$$

with unit determinant, the substitution is non-loxodromic when, and only when, $a + d$ is real; that is, when the characteristic equation

$$\begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = 0$$

is a *real* equation. The substitution is then hyperbolic, parabolic, or elliptic ‡ according as the two roots of this equation are real and distinct, coincident, or imaginary. Correspondingly, the projective substitutions are taken *exclusively* in this paper in homogeneous form,

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n' &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n, \end{aligned} \qquad (S) \tag{2}$$

* Presented to the Society, March 25, 1921.

† Fubini in his *Introduzione alla Teoria dei Gruppi discontinui e delle Funzioni automorfe*, pp. 77–8, has applied the above terms to those substitutions which are movements of

$$x_n^2 + \epsilon(x_1^2 + x_2^2 + \cdots + x_{n-1}^2) = 1 \qquad (\epsilon = \pm 1),$$

into itself. His aims and distinctions are very different from those of this paper.

‡ An application of the terms elliptic and hyperbolic differing slightly from the above usage of Klein is also to be found in the literature. Cf. Veblen's *Projective Geometry*, vol. 2, pp. 5, 171, 248.

with determinant Δ equal to $+1$. I call a substitution non-loxodromic, if, when so written, it possesses a real characteristic equation,

$$(3) \qquad \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0.$$

The characteristic roots are accordingly real or pairs of conjugate imaginaries.

It is well known that there are just two classes of groups (1) in a single variable which consist of non-loxodromic substitutions without a common pole. Either all the substitutions of the group leave invariant a circle in the complex x-plane and the group is called a "group with a principal circle" (Hauptkreis Gruppe), or the group consists entirely of elliptic substitutions. In the former case the group can be so transformed that the circle becomes the real axis. We have then a canonical form for the group in which its substitutions have real coefficients and unit-determinant, and, conversely, any group of this character is non-loxodromic. A second canonical form is

$$(3') \qquad x' = \frac{(a+bi)x + (c+di)}{(c-di)x + (a-bi)} \qquad (a^2+b^2-c^2-d^2=1),$$

which leaves the unit circle invariant. The second class of non-loxodromic groups—the elliptic groups—can be put into a canonical form in which the substitutions are all of form

$$(4) \qquad x' = \frac{(a+bi)x + (c+di)}{-(c-di)x + (a-bi)} \qquad (a^2+b^2+c^2+d^2=1),$$

and the group can be interpreted as a group of rotations of a sphere into itself.

When the number of homogeneous variables is greater than 2, there are, of course, more varieties of non-loxodromic substitutions to be distinguished than the familiar hyperbolic, elliptic and parabolic types, and it was perhaps scarcely to be expected that the result summarized in the preceding paragraph would have its exact analogue. Nevertheless an exactly similar result is obtained. There are still just two types of non-loxodromic groups. A limitation upon this statement (perhaps unnecessary) is introduced by the proof which requires that the non-loxodromic group should contain at least one substitution having a characteristic equation with n distinct roots. Corresponding to the hypothesis of the preceding paragraph that the substitutions (1) should not have a common pole, we have now the requirement that our non-loxodromic group of substitutions (2) in n homogeneous variables should not leave invariant a linear space of less than $n-1$ (complex) dimensions; in other words, should not transform into itself a hyperplane nor a space common to two or more hyperplanes in our projective space S_{n-1}.

One normal form obtained for the first class of non-loxodromic groups is that in which all substitutions (2) of the group have real coefficients and a determinant equal to $+1$. In this normal form the group obviously leaves invariant the real space embedded in the complex space S_{n-1} of the n homogeneous complex coördinates which is of $2n-2$ real dimensions. Any projective $(n-1)$-dimensional image of this real space within the complex space has been called a "chain" (von Staudt, Pieri), and every non-loxodromic group of our first class accordingly leaves some chain invariant. If $n-1$ is even, all projective substitutions which leave a chain invariant form a non-loxodromic group of the first class. When $n-1$ is odd, this is not the case. The most inclusive non-loxodromic group which leaves the chain invariant is then a subgroup of the set of all substitutions (2) which leave the chain invariant.*

The second class of non-loxodromic groups is novel and interesting. It occurs only when n is an even number $2m$. Any group of this class can be so transformed that all its substitutions have simultaneously the canonical form

$$
\begin{aligned}
x'_{2i-1} &= (a_{2i-1,\,1}\,x_1 + a_{2i-1,\,2}\,x_2) \\
&\qquad + \cdots + (a_{2i-1,\,2m-1}\,x_{2m-1} + a_{2i-1,\,2m}\,x_{2m}), \\
x'_{2i} &= (-\bar{a}_{2i-1,\,2}\,x_1 + \bar{a}_{2i-1,\,1}\,x_2) \\
&\qquad + \cdots + (-\bar{a}_{2i-1,\,2m}\,x_{2m-1} + \bar{a}_{2i-1,\,2m-1}\,x_{2m}) \\
&\qquad\qquad (i = 1, \cdots, m),
\end{aligned} \tag{5}
$$

where $\bar{a}_{ij}$ denotes the conjugate of a_{ij}. Thus the determinant of the substitution consists of m^2 blocks of form

$$
\begin{vmatrix} a_{ij} & a_{i,\,j+1} \\ -\bar{a}_{i,\,j+1} & \bar{a}_{ij} \end{vmatrix} \qquad (\text{odd } i \text{ and } j), \tag{6}
$$

similar to the block of coefficients of (4). It is shown that when a determinant has this structure, it is necessarily positive or zero.†

For any (non-singular) substitution of the second class the elementary divisors of (3) occur only in conjugate pairs $(\lambda - \lambda_i)^s$, $(\lambda - \bar{\lambda}_i)^s$ of equal degree, this statement applying as well for real roots λ_i which furnish pairs of equal elementary divisors. Conversely, a non-singular substitution which fulfills this requirement is one of the second class and may be transformed into the normal form (5). Such substitutions will be called *elliptic* (*restrictedly* elliptic, when the characteristic roots are m distinct pairs of conjugate imag-

* J. W. Young in these Transactions, vol. 11 (1910), p. 280, investigates the necessary and sufficient condition that a collineation ($n = 3$) shall leave a two-dimensional chain invariant.

† This result has been known for the special case in which all the elements of the determinant are real. Cf. Kowalewski, *Einführung in die Determinantentheorie*, pp. 156–158.

inaries). The characteristic property of an elliptic substitution in its normal form (5) is that it converts any pair of points

$$P = (x_1, x_2; \cdots; x_{2m-1}, x_{2m}), \qquad Q = (-\bar{x}_2, \bar{x}_1; \cdots; -\bar{x}_{2m}, \bar{x}_{2m-1})$$

into another pair of points whose coördinates are similarly related, any substitution having this property being necessarily of form (5). The invariant points of any substitution (5) occur in such pairs; in particular, the $2m$ invariant points of a restrictedly elliptic substitution, taken in normal form, consist of m distinct pairs.

It will be noticed that in (5) the $2m$ equations of the substitution are associated in pairs. In its normal form the most general group of the second or elliptic type comprises *all* non-singular elliptic substitutions (5) with the *same* pairing of the coördinates.

1. Derivation of two standard forms for non-loxodromic groups

As remarked in the introduction, we will suppose our projective substitutions to be written throughout in homogeneous form (2) with determinant

$$\Delta \equiv \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \tag{7}$$

equal to $+1$. Obviously the necessary and sufficient condition that a substitution (2) should be non-loxodromic by the above definition is that the sum of all the principal minors of each order $r \leqq n$ should be real.

When the roots of the characteristic equation (3) are all distinct, the substitution can be transformed through an appropriate substitution T into the canonical form

$$x_i' = \lambda_i x_i \qquad (i = 1, \cdots, n) \qquad (S_1). \tag{8}$$

We will treat only groups of non-loxodromic substitutions which contain at least one such substitution S_1, and we will suppose that the group has been so transformed that S_1 has the form (8). Further, it will be assumed that the subscripts i have been so assigned in (8) that for $i = 1, 2, \cdots, 2m \leqq n$ the λ_i are successively m pairs of conjugate imaginaries, while the remaining λ_i are real.

Consider now any non-loxodromic substitution S, given by (2), which combines with S_1 so as to generate a non-loxodromic group. The sum of principal minors $a_{11} + a_{22} + \cdots + a_{nn}$ is real, likewise the corresponding sum

$$\lambda_1^k a_{11} + \lambda_2^k a_{22} + \cdots + \lambda_n^k a_{nn} \equiv r_k$$

for SS_1^k. Denoting these real sums by r_k and taking $k = 0, 1, \cdots, n-1$, we have a system of n equations linear in the a_{ii}. The determinant of the

system is the product $\Pi(\lambda_i - \lambda_j)$ which cannot vanish in consequence of our hypothesis of the distinctness of the roots of S_1. There is therefore a unique solution for the a_{ii} in terms of the λ_i, r_i. The first $2m$ coefficients λ_i^k in each equation are successively pairs of conjugate imaginaries while the remaining ones are real. If, therefore, we exchange each of the first m odd terms on the left side of each equation with the following even term, we have a system of equations which is the conjugate of the preceding system, a_{ii} and $a_{i+1,\,i+1}$ being interchanged for odd $i < 2m$. Consequently a_{ii} and $a_{i+1,\,i+1}$ are conjugate imaginaries for odd values of $i < 2m$, while the a_{ii} are real for $i > 2m$.

Consider next the combination of (2) with any substitution

$$x_i' = b_{i1}x_1 + b_{i2}x_2 + \cdots + b_{in}x_n \qquad (S') \tag{9}$$

of the same non-loxodromic group. The product $S'S$ is a substitution

$$x_i' = B_{i1}x_1 + B_{i2}x_2 + \cdots + B_{in}x_n \qquad (i = 1, \cdots, n),$$

in which

$$B_{ij} = b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{in}a_{nj}. \tag{10}$$

Suppose first $m \geqq 1$. Then B_{11}, B_{22} are conjugate imaginaries, and the same is true of the corresponding coefficients

$$\begin{aligned} B_{11}^{(k)} &= \lambda_1^k b_{11}a_{11} + \lambda_2^k b_{12}a_{21} + \cdots + \lambda_n^k b_{1n}a_{n1}, \\ B_{22}^{(k)} &= \lambda_1^k b_{21}a_{12} + \lambda_2^k b_{22}a_{22} + \cdots + \lambda_n^k b_{2n}a_{n2} \end{aligned} \tag{11}$$

in $S'S_1^kS$. Take $k = 0, 1, \cdots, n-1$. We have then in (11) two systems of equations linear in the $b_{1i}a_{i1}$ and $b_{2i}a_{i2}$ respectively with a common determinant $\Pi(\lambda_i - \lambda_j)$ not zero. In the two systems corresponding coefficients λ_i^k are the same, the first m pairs being conjugate imaginaries and the remaining ones real, while the corresponding left members are conjugate imaginaries. We may therefore regard the two systems of equations as conjugate imaginary systems in which the $b_{1i}a_{i1}$ and $b_{2i}a_{i2}$ correspond to one another for $i > 2m$, while for odd $i < 2m$ they correspond respectively to $b_{2,\,i+1}a_{i+1,\,2}$ and $b_{1,\,i+1}a_{i+1,\,1}$ respectively. The solution of the two systems therefore gives

$$b_{1i}a_{i1} = \bar{b}_{2,\,i+1}\bar{a}_{i+1,\,2} \qquad (\text{odd } i < 2m), \tag{12}$$

$$b_{2i}a_{i2} = \bar{b}_{1,\,i+1}\bar{a}_{i+1,\,1} \qquad (\text{odd } i < 2m), \tag{13}$$

$$b_{1i}a_{i1} = \bar{b}_{2i}\bar{a}_{i2} \qquad (i > 2m). \tag{14}$$

We will now make the *Fundamental Hypothesis* that for each value of $i \leqq n$ there is in the non-loxodromic group at least one substitution (2) for which a_{1i} is not zero and also one for which a_{i1} is not, leaving for later con-

sideration the geometrical significance of this hypothesis. Choose then two substitutions S, S' for which $b_{21}\,a_{12} \neq 0$. Then by (13) b_{12} and a_{21} are not zero. Taking next S' identical with S, we see by (13) that $a_{12}\,a_{21}$ is its own conjugate and hence is real. A very different outcome is obtained according as we suppose that this product is positive or negative. The two corresponding cases we will refer to as *case* 1 *and case* 2 *respectively.*

To normalize the groups in these two cases it will suffice to make a simple change of variable

$$x_i = \rho_i\, y_i, \qquad x_i' = \rho_i\, y_i' \qquad (i = 1, \cdots, n) \tag{15}$$

with a proper determination of the ρ_i, restoring for convenience the letter x in place of y after the transformation. This change does not affect our canonical substitution S_1 nor the values of the a_{ii} in S, while the new value of a_{ij} differs from the old value by the factor ρ_j/ρ_i. Also the values of the principal minors of (7) and the characteristic equation (3) remain unaltered. We will so choose ρ_1, ρ_2 that in the substitution S of the preceding paragraph, for which $a_{12}\,a_{21} \neq 0$, we will have $a_{12} = \bar{a}_{21}$ in case 1 and $a_{12} = -\bar{a}_{21}$ in case 2. It follows then from (13) that we have simultaneously $b_{12} = \bar{b}_{21}$ in case 1, $b_{12} = -\bar{b}_{21}$ in case 2 for every substitution S' of the group.

Next take any odd value of i between 1 and $2m$ and choose two substitutions S, S' for which $b_{1i}\,a_{i1}$ is not zero. Then in consequence of (12) $a_{i+1,\,2}$ and $b_{2,\,i+1}$ are not zero. By properly choosing ρ_{i+1}/ρ_i we may make a_{i1} and $a_{i+1,\,2}$ in S the conjugates of one another. If then S' denotes any substitution whatsoever in the group (including S itself), it follows from (12) that for all substitutions S' we have simultaneously $b_{1i} = \bar{b}_{2,\,i+1}$ for odd $i < 2m$. Taking an S' for which $b_{1i} \neq 0$, we see again from (12) that for every substitution S of the group we have $a_{i1} = \bar{a}_{i+1,\,2}$ for odd $i < 2m$.

Consider next the two coefficients in $S'S_1^kS$

$$B_{12}^{(k)} = \lambda_1^k\, b_{11}\, a_{12} + \lambda_2^k\, b_{12}\, a_{22} + \cdots + \lambda_n^k\, b_{1n}\, a_{n2},$$
$$B_{21}^{(k)} = \lambda_1^k\, b_{21}\, a_{11} + \lambda_2^k\, b_{22}\, a_{21} + \cdots + \lambda_n^k\, b_{2n}\, a_{n1}.$$

These are conjugate imaginaries in case 1 and negative conjugates in case 2. If we set $k = 0, 1, \cdots, n-1$, we have two systems of equations linear in $b_{1i}\,a_{i2}$ and $b_{2i}\,a_{i1}$ respectively. Reasoning in exactly the same manner as above for equations (11), we obtain the following relations between any two substitutions of our group:

$$\begin{aligned} b_{1i}\,a_{i2} &= \bar{b}_{2,\,i+1}\,\bar{a}_{i+1,\,1} && \text{(case 1)} \\ b_{1i}\,a_{i2} &= -\bar{b}_{2,\,i+1}\,\bar{a}_{i+1,\,1} && \text{(case 2)} \end{aligned} \qquad (\text{odd } i < 2m) \tag{16}$$

$$\begin{aligned} b_{2i}\,a_{i1} &= \bar{b}_{1,\,i+1}\,\bar{a}_{i+1,\,2} && \text{(case 1)} \\ b_{2i}\,a_{i1} &= -\bar{b}_{1,\,i+1}\,\bar{a}_{i+1,\,2} && \text{(case 2)} \end{aligned} \qquad (\text{odd } i < 2m) \tag{17}$$

$$\begin{aligned} b_{1i}\,a_{i2} &= \bar{b}_{2i}\,\bar{a}_{i1} && \text{(case 1)} \\ b_{1i}\,a_{i2} &= -\bar{b}_{2i}\,\bar{a}_{i1} && \text{(case 2)} \end{aligned} \qquad (i > 2m) \tag{18}$$

For odd $i < 2m$ our normalization has already made $a_{i1} = \bar{a}_{i+1,\,2}$, $\bar{b}_{1i} = b_{2,\,i+1}$. If, therefore, we select for (17) a substitution S in which the former pair is not zero and for (16) a substitution S' in which the latter pair is not, we get from these equations $b_{2i} = \bar{b}_{1,\,i+1}$, $a_{i2} = \bar{a}_{i+1,\,1}$ in case 1, and $b_{2i} = -\bar{b}_{1,\,i+1}$, $a_{i2} = -\bar{a}_{i+1,\,1}$ in case 2. These results apply to any substitutions S', S in our group. Lastly, if we take a pair of substitutions with $b_{1i}\,a_{i1} \neq 0$ and divide (18) by (14), we get

$$\frac{a_{i2}}{a_{i1}} = \frac{\bar{a}_{i1}}{\bar{a}_{i2}} \qquad (i > 2m, \text{ case 1}), \tag{19}$$

while in case 2 these two ratios must be the negative of each other. But the latter is impossible, since a quantity cannot equal the negative of the conjugate of its reciprocal. We thus reach the important conclusion that *case 2 is possible only when* $i = 2m$.

In case 1 equation (19) makes $|a_{i1}| = |a_{i2}|$. We have not yet disposed of ρ_i for $i > 2m$, and we will now so take it that the arguments of a_{i1} and a_{i2} in the particular S under consideration shall be the negative of each other. Then we have $\bar{a}_{i1} = a_{i2}$, whence it follows by (18) that in case 1 *we have* $b_{2i} = \bar{b}_{1i}$ $(i > 2m)$ *for every substitution* S' *in the group.* By taking next an S' with $b_{1i} \neq 0$, it follows again from (18) that for every substitution S in the group we have $a_{i2} = \bar{a}_{i1}$ $(i > 2m)$ in case 1.

We have thus completely determined the character of the first two rows and columns of the coefficients (7) of every substitution in our non-loxodromic group. The character of the remaining rows and columns may now be rapidly fixed as follows. We know that for odd $j < 2m$ the coefficients

$$\begin{aligned} B^{(k)}_{1j} &= \lambda_1^k\, b_{11}\, a_{1j} + \lambda_2^k\, b_{12}\, a_{2j} + \cdots + \lambda_n^k\, b_{1n}\, a_{nj}, \\ B^{(k)}_{2,\,j+1} &= \lambda_1^k\, b_{21}\, a_{1,\,j+1} + \lambda_2^k\, b_{22}\, a_{2,\,j+1} + \cdots + \lambda_n^k\, b_{2n}\, a_{n,\,j+1} \end{aligned} \tag{20}$$

in $S'S_1^kS$ are conjugate imaginary in both cases 1 and 2. Putting $k = 0, 1, \cdots, n-1$, we have two systems of equations for the n quantities $b_{1i}\,a_{ij}$ and $b_{2i}\,a_{i,\,j+1}$ $(i = 1, \cdots, n)$ respectively. Since λ_i^k, λ_{i+1}^k are conjugate imaginaries for odd $i < 2m$ and λ_i is real for $i > 2m$, it follows in the same manner as before that we have the following pairs of conjugate imaginaries:

$$\begin{aligned} &b_{1i}\,a_{ij} \text{ and } b_{2,\,i+1}\,a_{i+1,\,j+1} \\ &b_{2i}\,a_{i,\,j+1} \text{ and } b_{1,\,i+1}\,a_{i+1,\,j} \end{aligned} \qquad (\text{odd } j < 2m \text{ and odd } i < 2m)$$

$$b_{1i}\,a_{ij} \text{ and } b_{2i}\,a_{i,\,j+1} \qquad (\text{odd } j < 2m \text{ and } i > 2m).$$

Now for odd $i < 2m$ we had already $\bar{b}_{1i} = b_{2,\,i+1}$ in both cases 1 and 2, $b_{1,\,i+1} = \bar{b}_{2i}$ in case 1 and $b_{1,\,i+1} = -\bar{b}_{21}$ in case 2; while for $i > 2m$ we had $\bar{b}_{1i} = b_{2i}$ in case 1. Since also by our *Fundamental Hypothesis* we could select S' so that a prescribed one of these quantities should not be zero, it follows that for every substitution S of our group we have the relations:

Case 1.

$$\begin{aligned} \bar{a}_{ij} &= a_{i+1,\,j+1} && (\text{odd } j < 2m;\ \text{odd } i < 2m) \\ \bar{a}_{i,\,j+1} &= a_{i+1,\,j} \\ \bar{a}_{ij} &= a_{i,\,j+1} && (\text{odd } j < 2m;\ \text{odd } i > 2m) \end{aligned}$$

Case 2.

$$\begin{aligned} \bar{a}_{ij} &= a_{i+1,\,j+1} \\ \bar{a}_{i,\,j+1} &= -a_{i+1,\,j+1} && (\text{odd } j \text{ and } i) \end{aligned}$$

The character of the coefficients has just been completely fixed for case 2 and for the first $2m$ columns of (7) in case 1. *In case* 2 *the determinant* (7) *consists of* m^2 *blocks of form* (6), *as stated in the introduction.*

It remains only to fix in case 1 the law of the coefficients subsequent to the $2m$th column of (7). Now it has already been found that $B_{1j}^{(k)}$ and $B_{2j}^{(k)}$ are conjugates for $j > 2m$, and their values will be given by (20) if the subscript $j+1$ in the second equation is replaced throughout by j. Taking then $k = 0, 1, \cdots, n-1$ in the two equations of (20) and arguing in the same manner as before, we obtain for odd $i < 2m$ the conjugate imaginary pairs $b_{1i}\,a_{ij}$ and $b_{2,\,i+1}\,a_{i+1,\,j}$, $b_{2i}\,a_{ij}$ and $b_{1,\,i+1}\,a_{i+1,\,j}$, while for $i > 2m$ we have as conjugates $b_{1i}\,a_{ij}$ and $b_{2i}\,a_{ij}$. For the given odd $i < m$ a substitution S' can be employed for which the conjugate elements b_{1i}, $b_{2,\,i+1}$ or b_{2i}, $b_{1,\,i+1}$ are not zero, and for $i > 2m$ a substitution for which the conjugates b_{1i}, b_{2i} are not. Hence, in every substitution S of our group, a_{ij} and $a_{i+1,\,j}$ are conjugates for odd $i < 2m$, $j > 2m$, while for $i > 2m$, $j > 2m$ every a_{ij} is self-conjugate and consequently real.

The completed result which has thus been reached in case 1 may now be summed up as follows. On the assumption that S_1 possesses distinct characteristic roots λ_i and at least one pair of roots which are conjugate imaginaries, a normal form for a non-loxodromic group satisfying our *Fundamental Hypothesis* has been obtained in which every substitution has the following structure. Its determinant (7) consists of four sections. The minor common to the first $2m$ rows and columns is made up of m^2 blocks of form

$$\begin{vmatrix} a_{ij} & a_{i,\,j+1} \\ \bar{a}_{i,\,j+1} & \bar{a}_{ij} \end{vmatrix} \qquad (\text{odd } i \text{ and } j),$$

and the minor common to the last $n-2m$ rows and columns contains only real elements. For $i < 2m$, $j > 2m$ the elements occur in conjugate pairs a_{ij}, $a_{i+1,\,j}$ (odd i), and for $i > 2m$, $j < 2m$ in conjugate pairs a_{ij}, $a_{i,\,j+1}$ (odd j). This normal form corresponds to the form (3′) given in the introduction for $m = 1$. The simple change of variable

$$(21) \qquad \begin{aligned} x_{2i-1} &= y_{2i-1} + iy_{2i}, \qquad x_{2i} = y_{2i-1} - iy_{2i} && (i < m), \\ x_i &= y_i && (i > 2m) \end{aligned}$$

throws the group at once into a second and simpler normal form in which the coefficients are all real.

When $m = 0$, the above derivation of the group fails, but the real normal form just indicated is still valid and can be rapidly derived as follows. By hypothesis, all the characteristic roots λ_i of S_1 are now real as well as distinct, and accordingly in every substitution (2) of the group all the a_{ii} are real. Since the left members of the system obtained by putting $k = 0, 1, \cdots, n-1$ in the first equation of (11) are real, all the $b_{1i}\, a_{i1}$ resulting from the solution of the system must be real. Further, by our *Fundamental Hypothesis*, for each value of i there is some S' for which b_{1i} is not zero, and in this particular S' by a change of variable (15) with the proper value of ρ_i/ρ_1 we may make b_{1i} $(i \neq 1)$ real. It follows then that in every substitution S of the group a_{i1} is real. By choosing next an S having $a_{i1} \neq 0$, it follows again from the reality of $b_{1i}\, a_{i1}$ that every substitution S' of the group has a real b_{1i}. In this wise all the elements of the first row and column of (7) may be made simultaneously real for all substitutions of the group. Turning next to the system of equations obtained by putting $k = 0, 1, \cdots, n-1$ in the first equation of (20), we argue in the same manner that every $b_{1i}\, a_{ij}$ is real and hence every a_{ij}. Thus by the simple change of variable (15) we can throw the group into a normal form in which all of its substitutions are real.

We turn now to consider the significance of our *Fundamental Hypothesis* that for each value of i there is some substitution of the group for which $a_{1i} \neq 0$. Make for a moment the contrary hypothesis and suppose that *for some fixed* j we have $a_{1j} = 0$ in every substitution of the group.* Then in (20) we have $B_{1j}^{(k)} = 0$ for this value of j. Taking $k = 0, 1, \cdots, n-1$, we have a homogeneous system of equations giving $b_{1i}\, a_{ij} = 0$ $(i = 1, \cdots, n)$ for *any* two substitutions S, S' of the group. Hence for any value of i we must have either $b_{1i} = 0$ for every substitution S' of the group or $a_{ij} = 0$ for every substitution S of the group. Suppose now that there are exactly r values of j for which a_{1j} is zero in every substitution of the group. Then for each of these r values of j and for any *other* value of i not one of these r values we must have $a_{ij} = 0$ in every substitution of the group. There will be no loss of generality if we assume that in the last r columns of (7) and in no others we have $a_{1j} = 0$, since this can be attained † by merely assigning the subscripts properly to our n variables x_i. Then all the elements of (7) common to these columns and to the first $n - r$ rows will be zero. Consequently the first $n - r$ equations of (2) contain only the variables $x_1, \cdots, x_{n-r}$. This reveals that the space of $r - 1$ (complex) dimensions (or a point, if $r = 1$) obtained

*If some of the a_{i_1} are zero for every substitution of the group, a parallel reasoning applies with like result.

†Unless $a_{11} = 0$ for every substitution, when we use the first and the last $r - 1$ columns in place of the last r columns.

by putting these coördinates equal to zero is carried over into itself by all substitutions of the group. Conversely, it is evident that if this is the case, the elements of (7) common to the last r columns and the first $n - r$ rows must be zero for all substitutions of the group. The significance of our *Fundamental Hypothesis* is therefore to forbid that all substitutions of the group should have such an invariant space when S_1 is taken in form (8).

The hyperplanes $x_i = 0$ are the only invariant hyperplanes of S_1 since the multipliers λ_i in (8) are by hypothesis all distinct. The linear spaces common to $n - r$ of these invariant hyperplanes are also invariant under S_1, and it is readily seen that they are the only spaces of $r - 1$ dimensions invariant under S_1 which are common to $n - r$ hyperplanes. For consider an invariant space which is defined by linearly independent hyperplanes

$$u_i \equiv \sum_{j=1}^{n} c_{ij} x_j = 0 \qquad (i = 1, \cdots, n - r).$$

If we take these planes as our coördinate planes along with r other independent hyperplanes, then S_1 in terms of the new u-coördinates will be specified by n equations of form (2) with the u_i in place of the x_i. In the corresponding determinant (7) the elements common to the last r columns and the first $n - r$ rows must be zero since the space common to the hyperplanes $u_i = 0$ $(i = 1, \cdots, n - r)$ is invariant. Accordingly there will be then $n - r$ linearly independent combinations of $u_1, \cdots, u_{n-r}$ corresponding to $n - r$ characteristic roots λ_i, and these, except for constant factors, must be identical with $n - r$ coördinates x_i which figure in (8). Hence the invariant space common to the hyperplanes $u_i = 0$ $(i = 1, \cdots, n - r)$ can also be specified by putting $n - r$ coördinates x_i equal to zero. We thus see that *our Fundamental Hypothesis is identical with the requirement that the substitutions of our group should not admit a common invariant hyperplane nor an invariant space common to two or more hyperplanes.* This requirement is, of course, unaltered when S_1 with the group is transformed out of its normal form. It corresponds to the condition mentioned in the introduction, that a non-loxodromic group of substitutions (1) should not have a common pole.

Up to the present point the development has involved only the reality of the characteristic equations (3) and in no wise the signs of the determinants (7). The conclusions reached therefore apply not only to groups of non-loxodromic substitutions but *likewise to non-singular groups of substitutions whose characteristic equations are real.* For the latter also, under the limitations specified, there are just two classes of groups, and these can be transformed into the two normal forms which have been given. Further, the totality of substitutions of either form, with or without the restriction of being non-loxodromic, form a group. This is evident in the case of substitutions of the

first kind since the totality of real, non-singular substitutions (2) form a group, and the same holds true if the requirement is added that the determinants of the substitutions shall be positive or that they shall be of value $+1$. In the next section it will be shown that the totality of non-singular substitutions of the second kind also form a group, in this case necessarily non-loxodromic.

2. On the nature of the two classes of non-loxodromic groups

A real projective substitution (2) of determinant Δ can be reduced to a form having the determinant unity by dividing each coefficient in (2) by $\Delta^{1/n}$. If Δ is positive, a real nth root may always be chosen so that the substitution remains in real form, but if Δ is negative, this is only possible when n is odd. A real projective substitution in space of $n-1$ (real or complex) dimensions is therefore necessarily non-loxodromic when the number of dimensions is even, but when it is odd, only if the determinant of the substitution is positive.

In accordance with § 1 the *general* non-loxodromic group of the first class, expressed in normal form, consists of all real substitutions (2) of determinant 1. Hence it comprises all non-loxodromic substitutions which leave invariant the real space of the n homogeneous variables which lies within the complex S_{n-1}. When $n-1$ is even, all real and non-singular projective substitutions of the n variables are included. If the group by transformation through any substitution is thrown out of normal form, the real space is converted into an invariant "chain" for the transformed group.

For the special case $n=2$ it is known that the poles of all hyperbolic and parabolic substitutions in a non-loxodromic group of the first class lie upon the invariant real axis or circle, while the poles of any elliptic substitution of the group lie symmetrically with respect to the real axis or circle. We can easily see that there is a corresponding generalization for greater values of n. Consider, for example, any non-loxodromic group of the first class which contains a substitution S whose characteristic roots are all real and *distinct*. Such a substitution will be called *hyperbolic* since it is the natural generalization of the hyperbolic substitution for $n=2$.* Let the group be transformed into its normal real form. Then the hyperbolic substitution becomes a real substitution (2) of unit determinant. To each of our n distinct characteristic roots λ_i belongs a linear form,

$$Y_i = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n,$$

which undergoes the substitution $Y_i' = \lambda_i Y_i$ in consequence of S. The well known equations for determining the coefficients c_i of the form, given in (23) below, show that the ratios of these coefficients are all real. If we equate

* It will presently appear that no hyperbolic substitution is included in a non-loxodromic group of our second class.

these n forms to zero, we obtain a configuration of n real and distinct hyperplanes, and the n vertices of this configuration are the poles of the hyperbolic substitution. The poles of the substitution are accordingly all real, a conclusion which applies simultaneously to all hyperbolic substitutions of the group. Hence *in a non-loxodromic substitution of the first class the poles of its hyperbolic substitutions all lie in the invariant chain of the group.*

Conversely, *if we have a set of hyperbolic substitutions whose poles lie in a common chain, they generate a non-loxodromic group.* For let the group be so transformed that the chain becomes the real coördinate space. The poles of any hyperbolic substitutions in the group are now real and by a real transformation of the group may be made to be the n-coördinate vertices. The substitution then has the real form $x_i' = \lambda_i x_i$ $(i = 1, \cdots, n)$ and hence must have had also a real form before the last transformation. Consequently all hyperbolic substitutions of the group have been made simultaneously real. Now by hypothesis our hyperbolic substitutions are of unit-determinant; they generate therefore a group consisting only of real substitutions of unit-determinant,—that is, a group of non-loxodromic substitutions.

Take now the other extreme case in which the characteristic roots of some particular substitution of a non-loxodromic group of the first class are distinct pairs of conjugate imaginaries. Such a substitution will be called *restrictedly elliptic* and is possible only when n is an even number $2m$. By transformation of the group the particular substitution can be thrown into the form (8) and be used as our substitution S_1 of the preceding section. Without change of S_1 the group was there brought into a first normal form and then was converted into the final real form by the transformation (21). Since the characteristic roots of our substitution are by hypothesis all distinct, the invariant points of (8) were the x-coördinate vertices. The transformation (21) converts the invariant point for which x_{2i-1} is the only non-zero coördinate into a point for which the y-coördinates are all zero except $y_{2i-1} = x_{2i-1}$, $y_{2i} = -ix_{2i-1}$, while the invariant point having x_{2i} as its only non-zero coördinate is converted into one for which all the y-coördinates are zero except $y_{2i-1} = x_{2i}$, $y_{2i} = ix_{2i}$. These two points are imaginary points which lie symmetrically with respect to the real space of the y-coördinates; i.e., the space for which the ratios of the $2m$ homogeneous coördinates are real.* Thus the $2m$ invariant

* If we think of a complex space S_k as a real space with $2k$ real orthogonal axes, two points lie symmetrically with respect to the linear space determined by r-coördinate axes if the corresponding r-coördinates of the two points agree, while the remaining coördinates of one point differ only in sign from those of the other. Accordingly, in the case before us, if we interpret the n complex coördinates y_i as nonhomogeneous coördinates for a real space of $2n$ dimensions, the $n = 2m$ invariant points of our substitution are obviously m pairs of points symmetrically situated with respect to the linear space of n dimensions which consists of the points for which the y_i are *all* real. When, however, as above, the coördinates are thought of as projective, some one of the y_i equated to zero—let us say $y_n = 0$—specifies the region at

points of the substitution S_1 in its final real form are m pairs of points symmetrically situated with respect to this real space of $2m - 1$ dimensions embedded in the complex space of the coördinates having double this number of dimensions. This real space is the invariant real space of the normalized real group in which our substitution is contained. The other "restrictedly elliptic" substitutions of the group, whose characteristic roots are *distinct* pairs of conjugate imaginaries, have become real simultaneously with S_1. Now any two *real* forms into which a given substitution can be transformed are convertible the one into the other by a real linear change of coördinate variables. Such a change in our y-coördinates converts a pair of points, symmetrically situated with respect to the real space of the coördinate-ratios and having corresponding coördinates the conjugates of one another, into another like pair of coördinates. Since, furthermore, any restrictedly-elliptic substitution of the group could be made our S_1, it follows that the invariant points of all restrictedly elliptic substitutions of a normalized non-loxodromic group of the first species consist of pairs of points which lie symmetrically with respect to the real invariant space of the group. If the group is transformed through any substitution T, the real invariant space is converted into a chain, and the images of the invariant point-pairs of these substitutions are converted into pairs of points which may be said to lie symmetrically with respect to this chain.

Turn now to the groups of the second class which are more novel and therefore more interesting. We will first prove that any substitution whose determinant consists of m^2 blocks of form (6) has a real characteristic equation. To establish this it is necessary to show that in its determinant (7) the sum of all the principal minors of any given order r is real. It suffices to prove that the principal minors which are not real occur in conjugate pairs. With this in view let us

(I) exchange each odd row and column of (7) with the following even row and column, and

(II) change first the signs of all the elements in the even rows and then the signs of the elements in the even columns of Δ.

By these two changes every element of (7) is replaced by its conjugate. Since the determinant is unaltered in value by the two changes, it must be real, being identical with its conjugate. Any principal minor of (7) formed

---

infinity and the nonhomogeneous coördinates will be $n - 1$ ratios $y_1/y_n, \cdots, y_{n-1}/y_n$. The difficulty arising from the presence of pairs of points for which y_n equals zero may be removed by regarding such a pair as the limit of another pair with the same coördinates except y_n, which is replaced by a real number h approaching zero. If, then, for $h \neq 0$ the two points are symmetrically situated with respect to the real space for which the $(n - 1)$-coördinate ratios are real, the limiting points $(h = 0)$ will also be said to lie symmetrically with respect to this space.

by deleting with each odd row and column the following even row and column is real for the same reason. Any other principal minor in consequence of I is exchanged with another principal minor, but neither is changed in sign by II. Hence these two minors must be the conjugate of one another, and the substitution has accordingly a real characteristic equation, as stated.

It will be seen later that a non-zero determinant Δ having the structure (6) must be positive. If therefore the substitution is non-singular, its determinant can be made equal to $+1$ by dividing all its coefficients by a real $\sqrt[n]{\Delta}$. *All non-singular substitutions of form* (5) *are therefore non-loxodromic.*

Consider next the product of any two (non-singular or singular) substitutions S, S' of the second class with coefficients a_{ij} and b_{ij} respectively, and with the same pairing of equations in each, as indicated in (5). Since the coefficients of both obey the law (6), we find at once from (10) that $B_{i+1,\,j+1} = \overline{B}_{ij}$, $B_{i+1,\,j} = -\overline{B}_{i,\,j+1}$ for odd i and j. Consequently the product of the two substitutions has the same form.

Furthermore, the inverse of a non-singular substitution will also be of like form. In fact, since its determinant (7) is real, it is only necessary to see that the determinant of the cofactors A_{ij} of the elements of (7) consists of blocks of the form (6). Now we have just seen that the changes I and II together replace every element of (7) by its conjugate and therefore every A_{ij} by its conjugate. On the other hand, these changes obviously result in an interchange of A_{ij} and $A_{i+1,\,j+1}$ for odd i and j, also of $A_{i+1,\,j}$ and $A_{i,\,j+1}$, with a reversal of sign in the latter case. Hence we have $A_{i+1,\,j+1} = \overline{A}_{ij}$, $A_{i+1,\,j} = -\overline{A}_{i,\,j+1}$. Thus the inverse substitution has the same form, and we conclude that *any number of non-singular substitutions* (5) *of our second kind generate a non-loxodromic group.*

We now proceed to examine the nature of the substitutions of our second class. Let λ_i be a characteristic root of any substitution (2) of the second class. To this there corresponds a linear form

$$Y_1 = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n \qquad (n = 2m), \tag{22}$$

which in consequence of the substitution undergoes the transformation $Y_1' = \lambda_i Y_1$, and its coefficients satisfy the equations

$$\begin{aligned} &c_1(a_{11} - \lambda_i) + c_2 a_{21} + \cdots + c_n a_{n1} = 0, \\ &c_1 a_{12} + c_2(a_{22} - \lambda_i) + \cdots + c_n a_{n2} = 0, \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ &c_1 a_{1n} + c_2 a_{2n} + \cdots + c_n(a_{nn} - \lambda_i) = 0, \end{aligned} \tag{23}$$

the consistency of which is guaranteed by our characteristic equation. Suppose now every quantity which appears in these equations to be replaced by its conjugate. We get a system of equations for $\bar{c}_1, \cdots, \bar{c}_n$. Exchange next

each odd equation in (23) with the following even equation, then each odd column with the following even column, and finally change the signs of the n terms of the resulting even equations. Then by virtue of the fact that a_{ij}, $a_{i+1,\,j+1}$, also $a_{i+1,\,j}$ and $-a_{i,\,j+1}$ are conjugates for odd i and j, we get precisely the equations which say that

$$-Z_1 = (\bar{c}_2 x_1 - \bar{c}_1 x_2) + (\bar{c}_4 x_3 - \bar{c}_3 x_4) + \cdots + (\bar{c}_n x_{n-1} - \bar{c}_{n-1} x_n) \tag{24}$$

undergoes the transformation $Z_1' = \bar{\lambda}_i Z_1$ in consequence of our substitution (2). Thus the pair of linear forms which correspond to a simple pair of conjugate imaginary roots of the characteristic equation have the form (22) and (24). Moreover, the two forms cannot be linearly dependent even when λ_i is self-conjugate and real, for then we would have

$$\frac{\bar{c}_{2i}}{c_{2i-1}} = -\frac{\bar{c}_{2i-1}}{c_{2i}}, \quad \text{or} \quad c_{2i}\,\bar{c}_{2i} + c_{2i-1}\,\bar{c}_{2i-1} = 0 \quad (i = 1, \cdots, m),$$

which would make every $c_i = 0$ in (22). It follows that every real root of the characteristic equation must be at least a double root.

In case the elementary divisors of the characteristic equation are not all simple, there correspond s linear forms to each elementary divisor $(\lambda - \lambda_i)^s$ of the x_i which in consequence of the substitution undergo the transformation

$$\begin{aligned} Y_1' &= \lambda_i Y_1, \\ Y_2' &= Y_1 + \lambda_i Y_2, \\ Y_3' &= Y_2 + \lambda_i Y_3, \\ &\cdots\cdots\cdots \\ Y_s' &= Y_{s-1} + \lambda_i Y_s. \end{aligned} \tag{25}$$

Let us suppose that Y_1 is given by the right member of (22), and correspondingly Z_1 by (24). Put

$$Y_2 = d_1 x_1 + d_2 x_2 + \cdots + d_n x_n \qquad (n = 2m). \tag{26}$$

By use of (1) and (26) in the second equation of (25) we get

$$\begin{aligned} &d_1(a_{11} - \lambda_i) + d_2 a_{21} + \cdots + d_n a_{n1} = c_1, \\ &d_1 a_{12} + d_2(a_{22} - \lambda_i) + \cdots + d_n a_{n2} = c_2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &d_1 a_{1n} + d_2 a_{2n} + \cdots + d_n(a_{nn} - \lambda_i) = c_n. \end{aligned} \tag{27}$$

Replace every quantity which appears in these equations by its conjugate, exchange each odd equation with the following even equation, then exchange each odd column with the following even column, and finally change the signs of the terms on both sides of the even equations. We get a set of equations which state that the linear form

$$-Z_2' = (\bar{d}_2 x_1 - \bar{d}_1 x_2) + \cdots + (\bar{d}_n x_{n-1} - \bar{d}_{n-1} x_n) \tag{28}$$

undergoes the transformation

$$Z_2' = Z_1 + \bar{\lambda}_i Z_2$$

in consequence of our substitution. Proceeding next in the same way with the third equation of (25) and so on in turn with each of its succeeding equations, we find that to the s linear forms $Y_1, \cdots, Y_s$ there correspond also s other linear forms $Z_1, \cdots, Z_s$, which are related to them in exactly the same way as (24) to (22) and which undergo the transformation

$$\begin{aligned} Z_1' &= \bar{\lambda}_i Z_1, \\ Z_2' &= Z_1 + \bar{\lambda}_i Z_2, \\ &\cdots\cdots\cdots \\ Z_s' &= Z_{s-1} + \bar{\lambda}_i Z_s, \end{aligned} \tag{29}$$

in consequence of the substitution. The coexistence of (29) and (25) indicates that for each elementary divisor $(\lambda - \lambda_i)^s$ there is another divisor $(\lambda - \bar{\lambda}_i)^s$, and hence gives the following result:

THEOREM. *The elementary divisors of the characteristic equation of any non-singular substitution* (5) *occur only in conjugate pairs* $(\lambda - \lambda_i)^s$, $(\lambda - \bar{\lambda}_i)^s$ *of equal degree, this conclusion holding also for the real characteristic roots* λ_i *to which belong pairs of equal elementary divisors.*

Since the determinant of a substitution is the product of the characteristic roots, taken each with its proper multiplicity, the theorem just stated gives at once the following important corollary:

COROLLARY. *If a non-singular substitution has the form* (5), *its determinant is positive.*

Determinants with *real* elements satisfying the condition (6) have been previously encountered in some work of Voigt and have been called *Determinants of Voigt* by Kowalewski. The proofs* given to show that Voigt's determinant is positive or zero fail altogether to cover the more general determinants here considered in which the elements are complex. As the proof furnished above was indirect, I shall also give a simple and direct proof.

The non-negative character of the determinant is obvious when $m = 1$ since it is then $a_{11}\bar{a}_{11} + a_{12}\bar{a}_{12}$. The extension to any other value of m is obtained in the following manner by mathematical induction. Suppose it to have been established for any value of m that all determinants of order $2m - 2$ having the structure (6) are non-negative. Take then any de-

* Kowalewski, *Einführung in die Determinantentheorie*, § 69, p. 156. Baltzer's and Drude's proofs are both found in the Göttinger Nachrichten, 1887.

Drude and Kowalewski prove, when the elements of (7) are real, that Δ is the sum of the squares of two expressions. When the elements are complex, the natural extension would be to replace each square by the product of the expression and its conjugate. But the resulting sum no longer represents Δ.

terminant (7) of order $2m$, which obeys the same law, and multiply it by *any* minor Δ_{2m-2} of order $2m-2$ obtained by deleting from Δ any odd row and any odd column and the following even row and column. This minor has the structure (6), and there will be no loss of generality in supposing that it consist of the first $2m-2$ rows and columns of Δ since the deleted rows and columns may be moved to the bottom and to the right-hand border of Δ without violating the law of its formation. By a well known identity* we then have

$$\Delta_{2m-2}\Delta = A_{2m-1,\,2m-1}A_{2m,\,2m} - A_{2m-1,\,2m}A_{2m,\,2m-1}. \tag{30}$$

Now we have already observed that $A_{2m-1,\,2m-1}$ and $A_{2m,\,2m}$ are conjugates, while $A_{2m-1,\,2m}$ and $A_{2m,\,2m-1}$ are negative conjugates. Hence the right member of (30) is non-negative, being positive unless all four A_{ij} on the right of (30) are zero. If, therefore, a single one of the minors Δ_{2m-2} formed in the manner indicated is positive, it follows that Δ is positive or zero. On the other hand, if every minor Δ_{2m-2} is zero, every A_{ij} by (30) must be zero and hence also Δ itself. Consequently, *any determinant consisting of blocks of form* (6) *must be non-negative.* Hence the determinant of any non-singular substitution of form (5) is positive.

Any non-singular substitution in n variables which consists only of pairs of components of form (25) and (29) is a substitution of our second class since it can be thrown into form (5) by alternating the equations of (29) with those of (25). Thus, for example, the alternation of their first two equations gives a matrix of coefficients

$$\begin{Vmatrix} \lambda & 0 & 0 & 0 \\ 0 & \bar{\lambda} & 0 & 0 \\ 1 & 0 & \lambda & 0 \\ 0 & 1 & 0 & \bar{\lambda} \end{Vmatrix}$$

having the structure (6).

In establishing the existence of the linear forms Y_i, Z_i for any non-singular substitution of the second class, we implicitly reduced it to canonical form. If we regard our m pairs of equations of form (22) and (24) as a change of variable, the substitution in terms of the new variables consists of pairs of components (25) and (29). Now the transforming equations (22) and (24) are of exactly the same form as (5). Hence any substitution of the second class can be thrown into canonical form by transformation through an appropriately chosen transformation of the same class. Since all the transformations (5) form a group, this can be also done without destroying the form (5) of a non-loxodromic group in which it is contained.

An important property of substitutions (5) *of our second class is to convert a pair of points*

$$P = (x_1, x_2; \cdots; x_{2m-1}, x_{2m}), \qquad Q = (-\bar{x}_2, \bar{x}_1; \cdots; -\bar{x}_{2m}, \bar{x}_{2m-1})$$

* Muir's *Theory of Determinants*, § 90.

into another similar pair. By direct substitution in (5) it will be found at once that when the substitution converts P into $P' = (x'_{2i-1}, x'_{2i})$, $i = 1, \cdots, m$, it also converts Q into $Q' = (-\bar{x}'_{2i}, \bar{x}'_{2i-1})$. A pair of such points we will speak of as *conjugate* with respect to the substitution or group.

We will now prove, conversely, that *if any substitution* (2) *in our* $2m$ *variables has the property of carrying every pair of conjugate points into a pair of conjugate points, it must be one of our second class having the form* (5). Consider the first two equations of (2) which, for convenience, we will write

$$\begin{aligned} x'_1 &= a_1 x_1 + a_2 x_2 + \cdots + a_n x_n, \\ x'_2 &= b_1 x_1 + b_2 x_2 + \cdots + b_n x_n \qquad (n = 2m). \end{aligned}$$

The coefficients of either equation will, in general, be determined by the requirement that m pairs of conjugate points P_i, Q_i shall be transformed into conjugate pairs P'_i, Q'_i. We get thus the following sets of equations for the determination of our coefficients a_i, b_i:

$$\begin{aligned} x'_{i1} &= a_1 x_{i1} + a_2 x_{i2} + \cdots + a_{2m-1} x_{i,\,2m-1} + a_{2m} x_{i,\,2m}, \\ -\bar{x}'_{i2} &= -a_1 \bar{x}_{i2} + a_2 \bar{x}_{i1} + \cdots - a_{2m-1} \bar{x}_{i,\,2m} + a_{2m} \bar{x}_{i,\,2m-1} \\ &\qquad (i = 1, \cdots, m), \\ x'_{i2} &= b_1 x_{i1} + b_2 x_{i2} + \cdots + b_{2m-1} x_{i,\,2m-1} + b_{2m} x_{i,\,2m}, \\ \bar{x}'_{i1} &= -b_1 \bar{x}_{i2} + b_2 \bar{x}_{i1} + \cdots - b_{2m-1} \bar{x}_{i,\,2m} + b_{2m} \bar{x}_{i,\,2m-1} \\ &\qquad (i = 1, \cdots, m). \end{aligned}$$

The determinants of these two sets of equations are identical and must be positive or zero since they are composed of blocks of form (6). To exclude the value zero, it will be assumed that m pairs of conjugate points P, Q are taken which do not lie in a common hyperplane of our complex S_{n-1}. The numerators of the values obtained for a_1 and b_2 are determinants differing only in their first two columns which consist of m blocks

$$\begin{Vmatrix} x'_{i1} & x_{i2} \\ -\bar{x}'_{i2} & \bar{x}_{i1} \end{Vmatrix}, \qquad \begin{Vmatrix} x_{i1} & x'_{i2} \\ -\bar{x}_{i2} & \bar{x}'_{i1} \end{Vmatrix},$$

respectively, while the corresponding blocks in the numerators of a_2 and b_1 are

$$\begin{Vmatrix} x_{i1} & x'_{i1} \\ -\bar{x}_{i2} & -\bar{x}'_{i2} \end{Vmatrix}, \qquad \begin{Vmatrix} x'_{i2} & x_{i2} \\ \bar{x}'_{i1} & \bar{x}_{i1} \end{Vmatrix}.$$

The remaining pairs of columns are identical for all four numerators and are of structure (6). If now in the numerator of b_2 we exchange each odd row and column with the following even row and column and change first the signs of the even columns and then of the even rows, each element of the numerator determinant is converted into the conjugate of the corresponding

element in the numerator of a_1. Consequently $b_2 = \bar{a}_1$. A like manipulation of the numerator of b_1 gives $b_1 = -\bar{a}_2$. In the same way we obtain $b_{2i} = \bar{a}_{2i+1}$, $b_{2i+1} = -\bar{a}_{2i}$. Hence the first two equations of the substitution must have the structure (6). The reasoning applies to all its succeeding pairs of equations. *Thus the property of converting pairs of conjugate points into conjugate pairs completely characterizes the second class of substitutions.*

One immediate corollary of this result is that the invariant points of a substitution of the second class consist of pairs of conjugate points, for clearly if one point of a conjugate pair is invariant, the other must be. The geometrical interpretation of this is well known for $m = 1$. If, namely, complex coördinates are given to the points of a sphere in the proper way, the formula (4) represents a rotation of the sphere about the diameter passing through the two invariant points of the sphere, and the coördinates* of any two diametrically opposite points are z, $-1/\bar{z}$, or, in homogeneous notation, (z_1, z_2) and $(-\bar{z}_2, \bar{z}_1)$.

It has thus appeared that *our second class of groups generalizes the characteristic form and property of the familiar elliptic rotation groups for* $m = 1$. For $m > 1$ it accordingly seems fitting to apply the term *elliptic* to any substitutions of our second kind, that is, to the substitutions whose elementary divisors occur only in conjugate imaginary or duplicate real pairs. The general elliptic group will then consist of all non-singular substitutions of the second class *with the same pairing of the coördinates*, as indicated in equation (5). A subclass will comprise the restrictedly elliptic substitutions whose characteristic roots are *n distinct* pairs of conjugate imaginaries.

* Cf. Klein's *Vorlesungen über das Ikosaeder*, pp. 32–34.

UNIVERSITY OF WISCONSIN,
MADISON, WIS.

BIRATIONAL TRANSFORMATIONS SIMPLIFYING SINGULARITIES OF ALGEBRAIC CURVES*

BY

GILBERT AMES BLISS

In a preceding paper† I have commented on the proofs of the theorem that every irreducible algebraic curve can be transformed by a birational transformation into one having no singularities except double points with distinct tangents. There are many proofs, two of which seem to be especially interesting. The first is by Walker,‡ who has worked out in detail an alteration, suggested by Klein in 1894,§ of a proof originally devised by Bertini for the projective plane. In the second, by Hensel and Landsberg,¶ reasoning proposed by Kronecker in 1881|| is extended to apply to the proof of the theorem in the function-theoretic plane. These two proofs appeal to me as being the best ones which I know for the two cases, but both of them are lengthy and complicated when all details are taken into consideration. In the present paper I have extended once more the method of Kronecker so that it can be applied in both the function-theoretic and the projective planes, and have attained what I hope will be regarded as simpler proofs of the two corresponding theorems.

It should perhaps be said that Kronecker's original reduction of the singularities of an algebraic curve to ordinary double points was effective only for the finite part of the plane. The alteration of his method to include the points at infinity in the function-theoretic plane is a real extension, as one may see by a study of the proof of Hensel and Landsberg. Still more is required of the transformation if higher singularities are to be excluded from points on the curve in the infinite region as well as in the finite part of the projective plane.

In Section 1 below a set of properties of an algebraic function $y(x)$ defined by an irreducible algebraic equation $f(x, y) = 0$ is described, such that when

* Presented to the Society, April 28, 1923.

† *The reduction of singularities of plane curves by birational transformation*, Bulletin of the American Mathematical Society, vol. 29 (1923), pp. 161–183.

‡ *On the resolution of higher singularities of algebraic curves into ordinary nodes*, Dissertation, Chicago, 1906.

§ See a footnote to the paper by Bertini, *Trasformazione di una curva algebrica in un'altra con soli punti doppi*, Mathematische Annalen, vol. 44 (1894), p. 158.

¶ *Theorie der algebraischen Funktionen*, p. 402, § 2; see also Hensel, *Encyclopädie der mathematischen Wissenschaften*, II C 5, § 25.

|| Journal für die reine und angewandte Mathematik, vol. 91 (1881), p. 301.

the function $y(x)$ has these properties the corresponding algebraic curve $f = 0$ has no singularities in the projective plane except double points with distinct tangents. For the simplification of the singularities of an arbitrary algebraic curve $f(x, y) = 0$ the method followed in the succeeding sections consists in the construction of a pair of rational functions $\xi(x, y), \eta(x, y)$, satisfying an irreducible equation $\phi(\xi, \eta) = 0$ birationally related to $f(x, y) = 0$, and having the properties described in Section 1. In Sections 2, 3, 4 the construction of ξ and η is described and their properties determined. The lemma in Section 2 is particularly effective in simplifying the proofs. In Section 5 the method of the preceding sections is applied in the function-theoretic plane, the general plan being that of Hensel and Landsberg, but the details quite different. In Section 6 the distinction between the theorems for the projective and the function-theoretic planes is explained in terms of homogeneous coördinates.

1. **A preliminary lemma.** Consider an equation $f(x, y) = 0$ in which f is an irreducible polynomial in x and y of degree n in y. Near a finite value $x = a$ the values of y which satisfy this equation are defined by a number of series in increasing integral powers of t of the form

$$y = bt^{\mu} + b' t^{\mu'} + \cdots, \tag{1}$$

where $t = (x - a)^{1/r}$, r is a positive integer, and the numbers $r, \mu, \mu', \cdots$ are relatively prime. Each such series defines r values of y when the r values of $(x - a)^{1/r}$ are substituted, and the sum of the integers r for the various series in powers of $(x - a)$ is equal to n. There are only a finite number of values $x = a$, the branch values, for which some of the corresponding exponents r are greater than unity. The polynomial $\Delta(x)$ is defined to be the product of the factors $(x - a)^{r-1}$ corresponding to the series (1) which have $r > 1$,

$$\Delta(x) = \Pi(x - a)^{r-1}.$$

Near $x = \infty$ the values of y are given by expansions (1) in which $t = (1/x)^{1/r}$.

LEMMA 1. *An irreducible algebraic curve $f(x, y) = 0$ of degree n in y has no singularities in the projective xy-plane except double points with distinct tangents, provided that the algebraic function $y(x)$ defined by the equation has the following properties:*

1. *at $x = \infty$ the n values of $y(x)$ are given by n expansions*

$$y = x\left(b_k + b_k' \frac{1}{x} + \cdots\right) \quad (k = 1, \cdots, n; b_k \neq b_l \text{ if } k \neq l) \tag{2}$$

and $y(x)$ has no other poles;

2. *the discriminant of $y(x)$ has the form*

$$D(x) = V(x)^2 \Delta(x)$$

where $V(x)$ is a polynomial with roots distinct from each other and from those of $\Delta(x)$.

To prove this lemma let $f(x, y)$ be first expanded in powers of x, y in the form

$$f(x, y) = (x, y)_0 + (x, y)_1 + \cdots + (x, y)_\nu,$$

where $(x, y)_k$ is homogeneous and of degree k in x and y. From the assumption that the series (2) satisfy $f = 0$ it follows readily that $(1, b_k)_\nu = 0$ $(k = 1, \cdots, n)$, and therefore $\nu \geqq n$. It is clear that $(x, y)_\nu$ is exactly of degree n in y since f itself has by hypothesis this property, and furthermore that $\nu = n$, since otherwise $f(x, y)$ would contain y^n multiplied by a polynomial in x, and the function $y(x)$ would necessarily have poles at finite values of x, contrary to the property 1. The intersections of $f = 0$ with the line at infinity are defined by the equation $(x, y)_\nu = 0$ and are evidently distinct. None of them is a multiple point of the curve.

At a finite value $x = a$ distinct from the roots of $D(x)$ every root b of $f(a, y)$ is simple and $f = 0$ has no multiple point at (a, b).

If the roots of $f(x, y)$ corresponding to a value x are represented by y_i $(i = 1, \cdots, n)$, the discriminant has the value

$$D(x) = V(x)^2 \Delta(x) = \prod_{i<j} (y_i - y_j)^2 \quad (i, j = 1, \cdots, n). \tag{3}$$

For a root $x = a$ of $\Delta(x)$ the n values of y are defined by expansions (1) in positive integral powers of $(x - a)^{1/r}$ which have the form

$$y = b + b'(x - a)^{1/r} + \cdots, \tag{4}$$

since $y(x)$ has no pole at a finite value of x. Each of these series furnishes $r(r - 1)/2$ differences $y_i - y_j$ of order at least $1/r$ in $(x - a)$. Since each factor $y_i - y_j$ occurs twice in $D(x)$ it follows that for each such series $D(x)$ contains at least the factor $(x - a)^{r-1}$, whatever the values of b and b' may be. Since, by property 2, $D(x)$ can contain no more powers of $(x - a)$ than are contained in $\Delta(x)$, it follows readily that for each series (4) corresponding to $x = a$ the coefficient b' must be different from zero, and that the coefficients b for the different series corresponding to $x = a$ must be distinct. The value b is therefore exactly an r-tuple root of $f(a, y)$, and $(y - b)^r$ is the lowest power of $(y - b)$ alone actually occurring in the expansion

$$f(x, y) = (x - a, y - b)_1 + \cdots + (x - a, y - b)_n \tag{5}$$

of $f(x, y)$ at (a, b). It may also be verified that the expansion (4) can satisfy $f = 0$ only if the term in $(x - a)$ is present in $(x - a, y - b)_1$. Hence none of the points (a, b) corresponding to a root $x = a$ of $\Delta(x)$ is a singular point of the curve $f = 0$.

At a zero $x = a$ of $V(x)$ the expansions (4) all have $r = 1$, since the roots of $V(x)$ and $\Delta(x)$ are distinct by the property 2 of the lemma. Since, by the property 2, $D(x)$ has exactly the factor $(x - a)^2$, it follows that one only of the differences $y_i - y_j$ has a simple zero at $x = a$, and that for it the expansions of y_i and y_j have the form

$$\begin{aligned} y_i &= b + c_1(x - a) + \cdots, \\ y_j &= b + c_2(x - a) + \cdots, \end{aligned} \tag{6}$$

with $c_1 \neq c_2$. The value $y = b$ is a double root of $f(a, y)$, and $(y - b)^2$ is the lowest power of $(y - b)$ alone actually occurring in the expansion (5). The series (6) can make (5) vanish identically only if the remaining linear term in $(x - a)$ is absent and if $(1, c_1)_2 = (1, c_2)_2 = 0$. Hence (a, b) is a double point of the curve $f = 0$ with distinct tangents. The other roots b' of $f(a, y)$ are in this case simple, and the points (a, b') are simple points of the curve.

2. **Rational functions $\sigma(x, y)$ with simple poles.** When the series (1) and the value of x in terms of t are substituted in a rational function $\sigma(x, y)$ of the variables x, y, an expansion for σ is found of the form

$$\sigma(x, y) = ct^{\mu} + dt^{\mu+1} + \cdots. \tag{7}$$

The constant μ is a positive or negative integer, not necessarily the same as the integer μ in (1), and is called the order of $\sigma(x, y)$ at the expansion (1). On the Riemann surface T of $f(x, y) = 0$ the places P are in one-to-one correspondence with the expansions (1). Hence μ is also called the order of $\sigma(x, y)$ at P.

A symbol $\mathbf{D} = P_1^{\mu_1} \cdots P_s^{\mu_s}$, where the P's represent places on the Riemann surface and the μ's are positive or negative integers, is called a divisor. The sum $\mu_1 + \cdots + \mu_s = -q$ is the order of the divisor. A rational function $\sigma(x, y)$ is a multiple of the divisor $\mathbf{D}$ if its orders μ are $\geqq \mu_i$ at the places P_i $(i = 1, \cdots, s)$ and $\geqq 0$ elsewhere. One of the most important problems of the theory of algebraic functions* is the determination of the functions $\sigma(x, y)$ which are multiples of a divisor $\mathbf{D}$. It is a theorem that if $q > 2p - 2$, where p is the genus of the curve $f(x, y) = 0$, then the number of linearly independent multiples of $\mathbf{D}$ is exactly $\nu = q - p + 1$.†

* Hensel, *Encyclopädie der mathematischen Wissenschaften*, II C 5, p. 557.

† Consider two divisors Q, Q' with orders q, q', and whose product is in the differential class (W). Let $\{Q\}, \{Q'\}$ represent the numbers of linearly independent multiples of $1/Q$ and $1/Q'$, respectively. Then, by the Riemann-Roch theorem,

$$q + q' = 2p - 2, \qquad \{Q'\} = \{Q\} - q + p - 1.$$

See Hensel und Landsberg, *Theorie der algebraischen Funktionen*, pp. 301, 304, formulas (1*a*) and (II); and Hensel, loc. cit., pp. 557–8. If $q > 2p - 2$ then $q' = 2p - 2 - q < 0$, and the number $\{Q'\}$ of multiples of $1/Q'$ is zero. Hence in this case $\{Q\} = q - p + 1$. In the text above the $1/Q$ of this footnote is replaced by $\mathbf{D}$.

Consider now a divisor $\mathbf{D} = P_1^{-1} \cdots P_q^{-1}$ for which $q > 2p + 2$. The $\nu = q - p + 1$ linearly independent multiples $\sigma_1, \cdots, \sigma_\nu$ of this divisor are everywhere finite except possibly for simple poles at the places in $\mathbf{D}$. The coefficient c of the lowest power permitted in the expansion (7) for a multiple of $\mathbf{D}$ may be called the leading coefficient of σ at the corresponding place P, and for the functions $\sigma_1, \cdots, \sigma_\nu$ these coefficients will be denoted by $c_1, \cdots, c_\nu$.

LEMMA 2. *For the multiples $\sigma_1, \cdots, \sigma_\nu$ of a divisor $\mathbf{D} = P_1^{-1} \cdots P_q^{-1}$ $(q > ?p + 2)$ the matrix*

$$\begin{Vmatrix} c_1 & \cdots & c_\nu \\ c_1' & \cdots & c_\nu' \\ c_1'' & \cdots & c_\nu'' \\ c_1''' & \cdots & c_\nu''' \end{Vmatrix} \tag{8}$$

of leading coefficients at four arbitrarily selected places P, P', P'', P''' on the Riemann surface of $f(x, y) = 0$ has always rank 4.

Suppose that this were not true for a set of places P, P', P'', P'''. Then there would be at least $\nu - 3$ linearly independent functions of the form $v_1\sigma_1 + \cdots + v_\nu\sigma_\nu$ with $v_1, \cdots, v_\nu$ constants and with leading coefficients zero at P, P', P'', P'''. These functions would be multiples of the divisor $\mathbf{D}_1 = PP'P''P'''\mathbf{D}$ of order $4 - q$. But this is impossible, since $q - 4 > 2p - 2$ and the number of linearly independent multiples of $\mathbf{D}_1$ is therefore

$$q - 4 - p + 1 = \nu - 4.$$

COROLLARY 1. *The matrix corresponding to* (8) *for $k < 4$ places has always the rank k.*

If the place P' is replaced by P, and the second row in the matrix (8) by the second coefficients $d_1, \cdots, d_\nu$ of the expansion (7) for $\sigma_1, \cdots, \sigma_\nu$, then the lemma and its proof would still be valid. The following corollary is therefore also true:

COROLLARY 2. *At every place P the second coefficients $d_1, \cdots, d_\nu$ in the expansions of $\sigma_1, \cdots, \sigma_\nu$ are not all zero.*

3. **Properties of pairs of rational functions.** The properties described in this section for pairs of rational functions $\xi(x, y)$, $\eta(x, y)$ on the Riemann surface T of an irreducible algebraic equation $f(x, y) = 0$ are well known.* They are reproduced here, however, in a form which is especially adapted to the applications which are to be made in Section 4.

In the first place if $\xi(x, y)$ has q poles on the surface T, then for every constant α the function $\xi(x, y) - \alpha$ will have q zeros, and it is clear that at the q places of T where $\xi(x, y)$ has a fixed value the function $\eta(x, y)$ takes q values $\eta_1, \cdots, \eta_q$ which may or may not be distinct. To obtain analytic ex-

* Hensel und Landsberg, loc. cit., p. 247. Appel and Goursat, *Théorie des Fonctions algébriques*, p. 256.

pressions for $\eta_1, \cdots, \eta_q$ for values of ξ near $\xi = \alpha$, consider the expansions analogous to (7) for the functions $\xi(x, y) - \alpha$ and $\eta(x, y)$ near a zero of the former. These have the form

$$(9)\quad \xi - \alpha = ct^\rho + dt^{\rho+1} + \cdots, \quad \eta = c_1 t^\mu + d_1 t^{\mu+1} + \cdots \quad (c \neq 0, c_1 \neq 0).$$

The first equation can be solved for t as a series in $\tau = (\xi - \alpha)^{1/\rho}$ defining ρ distinct values of t for each ξ near α when the ρ values of τ are substituted. When this series is substituted in the second equation the result has the form

$$(10)\qquad \eta = \beta\tau^\mu + \beta'\tau^{\mu'} + \cdots.$$

The q values $\eta_1, \cdots, \eta_q$ of $\eta(x, y)$ corresponding to a value of ξ near $\xi = \alpha$ are defined by a number of equations (10), one for each of the zeros of $\xi(x, y) - \alpha$. The sum of the orders ρ of these is q. For values ξ near $\xi = \infty$ the only difference is that the parameter in (10) is $\tau = (1/\xi)^{1/\rho}$.

The symmetric functions of the values $\eta_1, \cdots, \eta_q$ corresponding to a given ξ are rational functions of ξ. For they are clearly representable by series in fractional powers of $(\xi - \alpha)$ or $1/\xi$ near $\xi = \alpha$ or $\xi = \infty$, and these series can have only integral exponents since the symmetric functions which they represent are single-valued in ξ. Hence the symmetric functions, being single-valued in ξ and having no singularities except poles, are rational in ξ. It follows now that the product $(\eta - \eta_1)(\eta - \eta_2) \cdots (\eta - \eta_q)$ is a polynomial of degree q in η with coefficients rational in ξ, and when cleared of fractions it becomes a polynomial $\phi(\xi, \eta)$ in both ξ and η. The equation $\phi(\xi, \eta) = 0$ of degree q in η is satisfied by all of the values of $\xi(x, y)$, $\eta(x, y)$ at places on T.

The polynomial $\phi(\xi, \eta)$ is either irreducible or else the power of a single irreducible factor. In fact, each irreducible factor of $\phi(\xi, \eta)$, being made zero by one at least of the expansions (10), must also be caused to vanish identically by (9) and all of its continuations on the surface T. But in that case the irreducible factor must be satisfied by all of the roots $\eta_1, \cdots, \eta_q$, and the factors of $\phi(\xi, \eta)$ must all be identical. The important case when $\phi(\xi, \eta)$ itself is irreducible will evidently occur if and only if there is at least one value ξ for which all of the values $\eta_1, \cdots, \eta_q$ are distinct.

When $\phi(\xi, \eta)$ is itself irreducible the equation $\phi(\xi, \eta) = 0$ is birationally related to $f(x, y) = 0$. For then the set of series (10) corresponding to $\xi = \alpha$ must define distinct values $\eta_1, \cdots, \eta_q$ for values of ξ near $\xi = \alpha$, and the series themselves must all be distinct. To each one of them corresponds therefore a unique place on the surface T, which is equivalent to saying that there corresponds to each place on the Riemann surface U of $\phi(\xi, \eta) = 0$ a unique place P on the surface T. The values x, y belonging to P are therefore single-valued functions on the surface U. They have at most poles, since as

functions of t they have only poles, and since t is expressible in positive integral powers of τ by solving the equation

$$\tau = t(c + dt + \cdots)^{1/\rho}$$

obtained from the first equation (9) and the equation $\xi - \alpha = \tau^\rho$. Under these circumstances x and y are necessarily rational functions of ξ, η, and the curves $f(x, y) = 0$ and $\phi(\xi, \eta) = 0$ are birationally related.

4. **Transformation to a curve with ordinary double points only in the projective plane.** In this section it is proposed to prove the following principal theorem of this paper:

An irreducible algebraic curve $f(x, y) = 0$ can always be transformed by a birational transformation into a second such curve $\phi(\xi, \eta) = 0$ having no singularities in the projective $\xi\eta$-plane except ordinary double points.

The method of proof is to construct, with the help of the functions $\sigma_1, \cdots, \sigma_\nu$ of Section 2, a pair of rational functions $\xi(x, y)$, $\eta(x, y)$ satisfying an algebraic equation $\phi(\xi, \eta) = 0$ birationally related to $f(x, y) = 0$ and having the properties 1 and 2 of Lemma 1.

Consider now a divisor $\mathbf{D} = P_1^{-1} \cdots P_q^{-1}$ with $q > 2p + 2$, as in Section 2, having the linearly independent multiples $\sigma_1, \cdots, \sigma_\nu$. A particular function

$$\xi(x, y) = u_1\sigma_1 + \cdots + u_\nu\sigma_\nu$$

can be chosen with the constants $u_1, \cdots, u_\nu$ fixed once for all so that at each of the places $P_1, \cdots, P_q$ the order of ξ is exactly -1 (Lemma 2, Corollary 1 for $k = 1$). For a second function of the family

$$\eta(x, y) = v_1\sigma_1 + \cdots + v_\nu\sigma_\nu \tag{11}$$

the expansions (10) corresponding to $\xi = \infty$ arise at the places $P_1, \cdots, P_q$, and are readily seen to have the form

$$\eta = \xi\left(\beta_k + \beta_k'\frac{1}{\xi} + \cdots\right) \qquad (k = 1, \cdots, q). \tag{12}$$

The constants β_k will be distinct provided that $v_1, \cdots, v_\nu$ do not satisfy a certain system L of linear equations (Lemma 2, Corollary 1 for $k = 2$). Under these circumstances the values $\eta_1, \cdots, \eta_q$ corresponding to a value of ξ near $\xi = \infty$ will be distinct, and ξ and η will satisfy an irreducible algebraic equation $\phi(\xi, \eta) = 0$ of degree q in η birationally related to $f(x, y) = 0$. Furthermore, since η is infinite only at the places $P_1, \cdots, P_q$ where ξ is also infinite, the only poles of the function $\eta(\xi)$ defined by $\phi = 0$ are those given by the q expansions (12), and these have now the properties prescribed in 1 of Lemma 1.

The discriminant of η with respect to ξ,

$$D(v_1, \cdots, v_\nu, \xi) = \prod_{i<j}(\eta_i - \eta_j)^2 \quad (i, j = 1, \cdots, q), \tag{13}$$

is a polynomial in $v_1, \cdots, v_\nu, \xi$. For it is symmetric in $\eta_1, \cdots, \eta_q$ and therefore rational in ξ, and it must be a polynomial in ξ since it is finite near every finite value of ξ. It may be represented as a product $W(v_1, \cdots, v_\nu, \xi)\Delta(\xi)$ where $\Delta(\xi)$ contains all factors of D containing only ξ, and where W has consequently no factor in ξ alone.

Near a finite value $\xi = \alpha$ the q values of η are defined by a number of series (10) which now have the form

$$\eta = \beta + \beta'(\xi - \alpha)^{1/\rho} + \cdots \tag{14}$$

since η has no poles except at the places $P_1, \cdots, P_q$ where ξ is also infinite. The ρ values of η defined by one of these equations provide $\rho(\rho - 1)/2$ factors $\eta_i - \eta_j$ in D, each repeated twice, and each of order at least $1/\rho$ in $(\xi - \alpha)$. Hence they contribute at least the factor $(\xi - \alpha)^{\rho-1}$ to $\Delta(\xi)$. If the coefficients $v_1, \cdots, v_\nu$ in (11) do not satisfy a certain system of linear equations, then the coefficients β' in the expansions (14) corresponding to $\xi = \alpha$ will all be different from zero (Lemma 2, Corollary 2), and the constants β will all be distinct (Lemma 2, Corollary 1 for $k = 2$). For a suitable special choice of the coefficients v, therefore, the discriminant D will contain no more powers of $(\xi - \alpha)$ than the product of the factors $(\xi - \alpha)^{\rho-1}$ corresponding to the different expansions (14) for $\xi = \alpha$, and it can consequently contain no more when the v's are indeterminates. It follows readily that $\Delta(\xi)$ is exactly the product $\Pi(\xi - \alpha)^{\rho-1}$ taken for all the expansions at branch values $\xi = \alpha$, i.e., the values for which the function $\xi(x, y) - \alpha$ has multiple roots.

The factor $W(v_1, \cdots, v_\nu, \xi)$ of the discriminant is irreducible or else a product of a number of irreducible factors which are also polynomials in $v_1, \cdots, v_\nu, \xi$. Near a value $\xi = \xi_0$ which is not a branch value of ξ the discriminant D, thought of as a polynomial in $v_1, \cdots, v_\nu$, is completely decomposable into a product of factors

$$\eta_i - \eta_j = v_1(\sigma_{1i} - \sigma_{1j}) + \cdots + v_\nu(\sigma_{\nu i} - \sigma_{\nu j}) \tag{15}$$

linear in $v_1, \cdots, v_\nu$ and with coefficients power series in $\xi - \xi_0$ with constant terms not all zero (Lemma 2, Corollary 1 for $k = 2$). Each irreducible polynomial factor of D must therefore be expressible near $\xi = \xi_0$ as the product of a number of distinct factors of the form (15), and each must occur at least twice in W since each linear factor (15) occurs twice in D. An irreducible factor of W could not occur more than twice, since otherwise there would have to be at least two linear factors $\eta_i - \eta_j$, $\eta_k - \eta_l$ having $(i, j) \neq (k, l)$ and with proportional coefficients. This would imply a relation

$$\eta_i - \eta_j = (\eta_k - \eta_l)g$$

holding identically in $v_1, \cdots, v_\nu$, where g is a series in $(\xi - \xi_0)$ with constant

term different from zero. But this would contradict the validity of Lemma 2 at $\xi = \xi_0$, or its Corollary 1 for $k = 3$ when only three of the integers i, j, k, l are distinct. It is clear then that each irreducible factor in W occurs exactly twice, and D has the form

$$D(v_1, \cdots, v_r, \xi) = V(v_1, \cdots, v_r, \xi)^2 \Delta(\xi),$$

where V is a product of distinct irreducible polynomials in $v_1, \cdots, v_r, \xi$ containing no factor in ξ alone.

Suppose now that the coefficients $v_1, \cdots, v_r$ have numerical values not satisfying the linear system of equations L. Then, as has been shown above, the functions $\xi(x, y)$, $\eta(x, y)$ satisfy an irreducible algebraic equation $\phi(\xi, \eta) = 0$ birationally related to $f(x, y) = 0$, and the only poles of the function $\eta(\xi)$ defined by $\phi(\xi, \eta) = 0$ are given by the expansions (12) which have the form prescribed in 1 of Lemma 1. If furthermore $v_1, \cdots, v_r$ are so chosen that the roots of the irreducible factors of $V(v_1, \cdots, v_r, \xi)$ are distinct from each other and from those of $\Delta(\xi)$, then D has the form prescribed in 2 of Lemma 1. Hence $\phi(\xi, \eta) = 0$ has in the projective plane only double points with distinct tangents.

5. **Transformation to a curve with ordinary double points only in the function-theoretic plane.** The projective xy-plane is the extension of the euclidean xy-plane which is obtained by introducing the homogeneous coördinates $x = x_1/x_3$, $y = x_2/x_3$. The function-theoretic plane, on the other hand, is obtained when the homogeneous coördinates $x = x_1/x_2$, $y = y_1/y_2$ are used. In the former case the infinite region is the line $x_3 = 0$. In the latter case it is the totality of points whose non-homogeneous coördinates have the form (∞, y) or (x, ∞), or whose homogeneous coördinates have $x_2 = 0$ or $y_2 = 0$.

An irreducible algebraic curve $f(x, y) = 0$ is said to have ordinary double points only in the function-theoretic xy-plane if its singularities in the finite part of the plane are ordinary double points, and if its points at infinity provide only simple points or double points with distinct tangents when they are transformed into the finite part of the plane by means of one or both of the transformations $x = 1/x'$ or $y = 1/y'$. It is possible to transform an irreducible algebraic curve $f(x, y) = 0$ into one with only ordinary double points in the function-theoretic plane by a birational transformation of the form

$$\xi = x, \qquad \eta = \eta(x, y),$$

as will be shown in the following paragraphs. The first step in the proof of this statement is the following lemma analogous to Lemma 1 of Section 1.

LEMMA 3. *An irreducible algebraic curve $f(x, y) = 0$ of degree n in y has no singularities except double points with distinct tangents in the function-theoretic*

plane, provided that the function $y(x)$ defined by the equation $f = 0$ has the following properties:

1. *at $x = \infty$ the function $y(x)$ has n distinct finite values*;

2. *the poles of $y(x)$ are simple and at values $x = a$ distinct from each other and from the branch values of x*;

3. *the discriminant of $y(x)$ has the property 3 of Lemma 1.*

It is clear from 1 that the points (∞, y) on the curve are simple when transformed into finite points by the transformation $x = 1/x'$, and from 2 the same is true of the points (x, ∞) when transformed by $y = 1/y'$.

The finite roots of $f(x_0, y)$ for a value $x = x_0$ over which $y(x)$ has a pole are all simple, by 2, and hence they provide only simple points of the curve

$$f(x, y) = 0.$$

The discriminant of $y(x)$ is by definition

$$D(x) = A^{2n-2} \prod_{i<j} (y_i - y_j)^2 \quad (i, j = 1, \cdots, n),$$

where A is the product of the linear factors $x - x_0$ belonging to values x_0 for which $y(x)$ has poles. By a repetition of the argument of Section 1 it is provable that $f(x, y) = 0$ has only double points with distinct tangents corresponding to values $x = a$ other than those discussed above, and the lemma is therefore proved.

To construct a rational function $\eta(x, y)$ whose irreducible algebraic equation $\phi(x, \eta) = 0$ is birationally related to $f(x, y) = 0$ and has the properties 1, 2, 3 of Lemma 3, one may start from a divisor $\mathbf{D} = P_1^{-1} \cdots P_q^{-1}$ with $q > 2p + 2$, as in Section 2, whose places P_i on T correspond to finite values of x distinct from each other and from the branch values. If the coefficients $v_1, \cdots, v_r$ in the expression

$$\eta(x, y) = v_1 \sigma_1 + \cdots + v_r \sigma_r$$

do not satisfy a certain system L_1 of linear equations, the function $\eta(x, y)$ will have distinct finite values at $x = \infty$ (Lemma 2, Corollary 1 for $k = 2$) and a pole of order one at each place P_i (Lemma 2, Corollary 1 for $k = 1$). Since under these circumstances the n values of η corresponding to a value x near $x = \infty$ are all distinct, it follows that $\eta(x, y)$ will satisfy an algebraic equation $\phi(x, \eta) = 0$ birationally related to $f(x, y) = 0$ and having the properties 1 and 2 of Lemma 3.

The discriminant of $\eta(x, y)$ with respect to x is by definition

$$D(v_1, \cdots, v_r, \xi) = A^{2n-2} \prod_{i<j} (\eta_i - \eta_j)^2 \quad (i, j = 1, \cdots, n),$$

where A is the product of the factors $(x - x_0)$ belonging to the places P_i of D, and where $\eta_1, \cdots, \eta_n$ are the n values of $\eta(x, y)$ for a given value x. This expression for D is symmetric in $\eta_1, \cdots, \eta_n$ and therefore rational in x. It can become infinite only at the values x_0 corresponding to the places P_i. At such a value x_0 only one of the values η_i has a simple pole. There are therefore only $n - 1$ of the factors $\eta_1 - \eta_j$ which have simple poles, and in D these are annulled by the factor A^{2n-2}. Hence D is finite for all finite values of x and is a polynomial in $v_1, \cdots, v_\rho, x$. It is expressible in the form

$$W(v_1, \cdots, v_\rho, x)\Delta(x)$$

where W contains no factor in x alone.

It may now be proved, exactly as in Section 4, that $\Delta(x)$ is the product $\Pi(x - a)^{r-1}$ formed for all the branch places on T, and that W is the square of a product $V(v_1, \cdots, v_\rho, x)$ of distinct irreducible factors. Hence if $v_1, \cdots, v_\rho$ are chosen not satisfying the system L_1 of linear equations, and so that the roots of V are distinct from each other and those of $\Delta(x)$, then $\phi(x, \eta) = 0$ will have the properties 1, 2, and 3 of Lemma 3, and the curve which it defines in the function-theoretic plane will have only double points with distinct tangents.

6. **Homogeneous coördinates.** The branches of an irreducible algebraic curve $f(x, y) = 0$ have for $x = \infty$ and $x = a$ the forms

$$x = t^{-r}, \qquad y = bt^{\mu} + b' t^{\mu'} + \cdots,$$
$$x = a + t^{r}, \qquad y = bt^{\mu} + b' t^{\mu'} + \cdots.$$

These, with their transforms after replacing t by a series $ct + dt^2 + \cdots$ $(c \neq 0)$, are included in the type

$$x = P(t), \qquad y = Q(t), \tag{16}$$

where P and Q are series in integral powers of t having at most a finite number of terms with negative exponents.

In terms of homogeneous coördinates $x = x_1/x_3$, $y = x_2/x_3$ the branch (16) takes the form

$$x_l = a_l + \alpha_l t + \cdots \qquad (l = 1, 2, 3) \tag{17}$$

where $(a_1, a_2, a_3) \neq (0, 0, 0)$. A branch of this character is said to be linear if the matrix

$$\begin{Vmatrix} a_1 & a_2 & a_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \end{Vmatrix}$$

is of rank 2. Two such branches with the same center $a_1 : a_2 : a_3$ have distinct tangents if the determinant

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_1' & \alpha_2' & \alpha_3' \end{vmatrix}$$

is different from zero, where α_1', α_2', α_3' are the coefficients of t in the second branch.

In the projective plane a curve $f(x, y) = 0$ has no singularities except double points with distinct tangents if all of its branches are linear, if the same center is never shared by more than two branches, and if every pair of branches with a common center has distinct tangents. It is readily provable that the branches (17) of a curve $f(x, y) = 0$ such as is described in Lemma 1 of Section 1 have these properties, and that the properties themselves are invariant under transformations of the parameter t and projective transformations

$$x_i = \alpha_{i1}\,\xi_1 + \alpha_{i2}\,\xi_2 + \alpha_{i3}\,\xi_3 \qquad (i = 1, 2, 3).$$

In terms of homogeneous coördinates $x = x_1/x_2$, $y = y_1/y_2$ the branches (16) take the form

$$\begin{aligned} x_i &= a_i + \alpha_i\, t + \cdots, \\ y_i &= b_i + \beta_i\, t + \cdots \end{aligned} \qquad (i = 1, 2).$$

Such a branch is said to be linear if the determinants

$$\begin{vmatrix} a_1 & \alpha_1 \\ a_2 & \alpha_2 \end{vmatrix}, \qquad \begin{vmatrix} b_1 & \beta_1 \\ b_2 & \beta_2 \end{vmatrix}$$

are not both zero. Two branches with the same center

$$(a_1 : a_2 | b_1 : b_2) = (a_1' : a_2' | b_1' : b_2')$$

have distinct tangents if the expression

$$\begin{vmatrix} a_1 & \alpha_1' \\ a_2 & \alpha_2' \end{vmatrix} \begin{vmatrix} b_1' & \beta_1 \\ b_2' & \beta_2 \end{vmatrix} - \begin{vmatrix} a_1' & \alpha_1 \\ a_2' & \alpha_2 \end{vmatrix} \begin{vmatrix} b_1 & \beta_1' \\ b_2 & \beta_2' \end{vmatrix}$$

is different from zero, where the primes designate the coefficients of the second branch.

A curve with no singularities except double points with distinct tangents in the function-theoretic plane is now defined as in the next to last paragraph above. The curves described in Lemma 3 of Section 5 have these properties, and the properties themselves are invariant under transformations of the parameter t and the transformations

$$\begin{aligned} x_i &= \alpha_{i1}\,\xi_1 + \alpha_{i2}\,\xi_2, \\ y_i &= \beta_{i1}\,\eta_1 + \beta_{i2}\,\eta_2 \end{aligned} \qquad (i = 1, 2).$$

THE UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

A SYMBOLIC THEORY OF FORMAL MODULAR COVARIANTS*

BY

OLIVE C. HAZLETT

PART I. INTRODUCTION

1. **Prologue.** Thus far, very little has been published on the general theory of formal modular invariants or covariants. Workers have, on the whole, obtained results for special, more or less isolated, cases; and although some beautiful and important general theorems have been proved, they are more or less unrelated. This is, of course, only natural in any division of knowledge in its formative state.

Nevertheless, no worker in the field could fail to be conscious of a certain uniformity common to the special cases that have been studied in detail; though (alas!) this uniformity usually appeared to be broken ruthlessly in the next case studied. This breaking of an apparent law signified, however, merely that we did not know these special cases with a sufficient thoroughness of illuminating detail, or were trying unwittingly to make the laws conform to certain standards, unconsciously preconceived. This latter handicap was laid on us naturally enough by our thorough knowledge of algebraic invariants and the fact that this newer kind of covariants is, in many ways, strikingly like the older, classic covariants, though so tantalisingly different.

Their similarity and their difference show themselves in the very beginning of the study: in the definitions, in the simplest examples. Perhaps the differences that first come to mind are those which are inherent in the fields of definition, which, in the case of classic covariants, is the field of reals or ordinary complex numbers and, in the case of modular covariants, is a Galois field, $GF[p^n]$, of order p^n. These differences are too obvious to mention in detail, but one who has studied the beautiful proofs given by the old masters of invariant theory has been forced to the conclusion that most of the proofs seemed to use the properties of a field of characteristic zero, not in some accidental manner, but rather in veriest necessity.

Growing from the surface differences between the two fields are two very important distinguishing characteristics of the two kinds of covariants. It

* Part II was presented to the Society, September 7, 1920; Part III, December 28, 1921; Parts IV and V, December 27, 1922.

The reading of the literature in connection with this paper was much facilitated by the purchase of books with a grant made by the American Association for the Advancement of Science and this help is herewith gratefully acknowledged.

is well known that an algebraic covariant is necessarily such that all of its terms have the same weight—i.e., it is isobaric; but a modular covariant is not necessarily isobaric. Nevertheless, although its terms are not in general of exactly the same weight, their weights can differ at most by multiples of $p^n - 1$. Moreover, if two classic invariants are identical in value for all marks of the field, they are necessarily identical in form, and conversely; whereas, if two formal modular invariants are identical for all marks of the field of definition, they are not necessarily identical in form. Nevertheless, although two such invariants can be of different degree and appear quite different in form at a casual glance, yet they necessarily have in common certain fundamental characteristics.* There are other differences that will occur to any worker in the field, but I think that I have mentioned the most refractory.

As indicated above, in spite of the great differences between the algebraic covariants and formal modular covariants, there are certain fundamental likenesses which are more easily sensed than they are analyzed. A feeling that there is some theory which underlies all the special cases, and yet which is comparatively simple, made the writer try to crystallize this theory into words.

In the spring of 1918, came the feeling that the theory of formal modular covariants of a binary form, f, for the field $GF[p^n]$, must be, in essentials, equivalent to the theory of simultaneous algebraic covariants of f and certain other forms obtained from f by replacing the coefficients of f by their respective p^nth powers. This is natural enough, since these powers of the coefficients of f are cogredient with the coefficients of f for the transformations of the group used, and this is the only way in which a formal modular covariant differs from an algebraic covariant of f. Then there appeared other indications that there is an intimate relation between formal modular covariants of f and algebraic covariants of a system of forms consisting of f and related forms. This paper is an attempt to put in systematic form the theorems which emerged when these eventually crystallized.

2. **Summary of literature.** Let

$$f(a; x) = a_0 x^m + a_1 x^{m-1} x_2 + \cdots + a_{m-1} x_1 x^{m-1} + a_m x^m$$

be a binary form and let G be a group of linear transformations

$$x_i = \xi_i x_1' + \eta_i x_2' \qquad (i = 1, 2) \tag{1}$$

whose coefficients, the ξ's and η's, are in the field F. If $I(a)$ be a polynomial in the a's which, under all transformations of the group, is transformed into $I(a')$ such that

$$I(a') = D^w I(a) \quad \text{(in the field)} \tag{2}$$

* Some of these characteristics are indicated in a paper by Hazlett, these Transactions, vol. 22 (1921), pp. 144–157, especially p. 145.

where

$$D = (\xi\eta) = \xi_1 \eta_2 - \xi_2 \eta_1,$$

then I is said to be an invariant of f under the group G. Similarly, if $C(a; x)$ is a polynomial in the a's and the x's which, under all transformations of the group G, is transformed into $C(a'; x')$ such that

$$C(a'; x') = D^w C(a; x) \quad \text{(in the field)}, \tag{3}$$

then C is said to be a covariant of f under the group G. Also, we speak of invariants and covariants of a system S of binary forms f_i.

When the field of definition, F, is the field of all reals or the field of all ordinary complex numbers, the covariants of S are the ordinary covariants of S of the classic invariant theory, with which are associated the names of Cayley, Sylvester, Hermite, Aronhold, Gordan, Clebsch and Hilbert.

When the field of definition, F, is the Galois field, $GF[p^n]$, of order p^n, the covariants of S are called modular covariants. Here the ξ's and η's are marks of the Galois field, $GF[p^n]$, of order p^n* defined by the prime p and an algebraic equation, $P(X) = 0$, of degree n; and thus (2) and (3) are congruences, reduced modulis p and $P(X)$. This means that, in (2) and (3), the left member is identical with the right member if we replace the p^nth power of ξ_i and of η_i by ξ_i and η_i respectively, in view of Galois' generalization of Fermat's theorem.† For a modular field there are two different types of covariants. If the coefficients of the forms (the a's) range over the marks of $GF[p^n]$, and if (2) and (3) are congruences which are true if Fermat's theorem is applied, not only to the ξ's and η's, but also to the a's, then the covariants are called simply *modular covariants*. If, on the other hand, the a's are independent variables, Fermat's theorem does not apply to them; and hence, in (2) and (3), the left member is understood to be identical with the right member without any reduction in the exponents of the a's. Such covariants are called *formal*

* Let p be any prime and let $P(X) = 0$ be any algebraic equation of degree n which has its coefficients integers reduced modulo p and which is irreducible, modulo p. Then, if we reduce any polynomial in X modulis p and $P(X)$, we obtain a polynomial of the form

$$M(X) = c_{n-1} X^{n-1} + c_{n-2} X^{n-2} + \cdots + c_1 X + c_0$$

where each of the c's is an integer of the set $0, 1, \cdots, p-1$. The totality of all polynomials congruent modulis p and $P(X)$ to the same $M(X)$ is said to form a class of residues. Since there is a class of residues for each set of values for the c's, there are p^n such classes. The totality of the classes of residues are closed under addition, subtraction, multiplication and division (provided the divisor is not zero) and so constitute a field containing p^n marks or elements. Any two such fields are the same for a given p and a given n. This is called a Galois field of order p^n and is denoted by $GF[p^n]$. Moreover, any finite field is simply isomorphic with a Galois field.

† If a is any mark of $GF[p^n]$, defined by p and $P(X)$, then $a^{p^n} \equiv a$ (modd p, $P(X)$). See any standard work on Galois fields, such as Dickson, *Linear Groups*, Teubner, 1901, p. 11; Serret, *Cours d'Algèbre Supérieure*, vol. 2, p. 180.

modular covariants. If the covariant is independent of the variables, it is called a *modular invariant* or a *formal modular invariant* according as the a's are marks of the field or independent variables.

Although Hurwitz introduced the notion of formal modular invariants in connection with the determination of the number of solutions of higher congruences,* Dickson discovered them independently four years later from a different point of view.† Practically all important results in the theory are due to Dickson.‡

Clearly, every formal modular covariant of a system S is a modular covariant of S, though not every modular covariant of S is also a formal modular covariant of S. For a modular invariant, I, of S is concerned solely with the values of I for the different sets of values of the a's in the field of definition, whereas a formal modular invariant of S is concerned not only with the values but also with the form of I. This statement has to be modified somewhat for covariants that are not invariants, since, in (2) and (3), the left member is the same as the right member in a purely formal sense as far as the x's are concerned, for both kinds of covariants. For example, if a is any one of the coefficients of f, then $a^{p^n} - a$ always has the value zero when a is a mark of $GF[p^n]$ and hence it is a modular invariant of f; but it is not a formal modular invariant, since it is changed in form if we interchange x_1 and x_2. To take a less trivial example,

$$q = (a + c)(b^2 + ac - 1)$$

is a modular invariant of $f = ax_1^2 + 2bx_1x_2 + cx_2^2$, mod 3, but is not a formal invariant, since $a + c$ goes into $-a - b + c$ under the transformation $x_1 = x_1' + x_2'$, $x_2 = x_2'$.

In 1909,§ Dickson studied modular invariants from a different point of view and introduced the notion of classes of forms. This enables one to see to the very heart of the theory and the finiteness theorem follows almost directly from the definition. Later,‖ he proved the finiteness theorem for modular covariants.

* *Ueber höhere Kongruenzen*, Archiv der Mathematik und Physik, ser. 3, vol. 5 (1903), pp. 17–27.

† *Invariants of binary forms under modular transformations*, these Transactions, vol. 8 (1907), pp. 205–232.

‡ Anyone wishing to gain familiarity with this beautiful theory should read the article by Hurwitz mentioned above and Dickson's papers, of which the most fundamental are I. *General theory of modular invariants*, these Transactions, vol. 10 (1909), pp. 123–158; II. *Proof of the finiteness of modular covariants*, ibid., vol. 14 (1913), pp. 299–310. Also one should read the brief but important paper by Miss Sanderson, *Formal modular invariants with application to binary modular covariants*, these Transactions, vol. 14 (1913), pp. 489–500. All results obtained up to 1914 are summarized in Dickson's *Madison Colloquium Lectures*, and all essential results published up to August, 1922, are summarized briefly in Chapter 19 of his *History of the Theory of Numbers*, vol. 3.

§ Dickson, I.

‖ Dickson, II.

There is, however, an intimate relation between modular invariants and formal modular invariants. For Miss Sanderson's theorem* tells us that, corresponding to any modular invariant, i, of a system S under any group G of linear transformations with coefficients in $GF[p^n]$, there is a formal invariant, I, under G such that $I \equiv i$ for all sets of values, in the field, of the coefficients of the system S. This enabled her to construct modular covariants of a system S of binary forms, f_i, from the modular invariants of another system S', consisting of the forms f_i and an additional linear form with coefficients x_2 and $-x_1$. Moreover, every modular covariant of S is a polynomial in the universal covariant, $L = x_1^{p^n} x_2 - x_1 x_2^{p^n}$, and the modular covariants obtained in the manner indicated by Miss Sanderson's theorem.† Since every algebraic covariant of S is obtained from the algebraic invariants of S' without the need of any additional covariant, analogous to L, this result shows one of the fundamental differences between algebraic and modular covariants.

In connection with this last remark, several minor results are of interest, as they point the way toward more general results. In 1920, it was shown that, for the field $GF[p^n]$ with $p \neq 2$, every modular covariant of a binary form, f, whose degree is not divisible by p, is expressible as a rational function of the universal modular covariants, L and Q, and of algebraic covariants of f. For formal modular covariants of f, there is a theorem which is more complicated in statement but similar in essence.‡ Then, in 1922, W. L. G. Williams announced that every formal modular seminvariant of f (aside from a power of a_0) is a polynomial in the algebraic protomorphs and in

$$\beta = a_1^p - a_1 a_0^{p-1}$$

for $GF[p]$ when the form f is such that binomial coefficients can be used.§ He also proved the analogous theorem for formal seminvariants of two or more binary forms. In this latter, it is interesting to note that, to a fundamental set of algebraic protomorphs, it is necessary to adjoin only one new seminvariant, namely, one of the same type as β but formed for any one of the forms f_i. It will be seen that these results indicate relations between modular covariants and algebraic covariants, in spite of their superficial differences.

3. **Summary of this paper.** We first explain a symbolic notation for formal modular covariants which is like that generally used for algebraic covariants in most respects, but is (necessarily) different from the latter in one essential.

* These Transactions, vol. 14 (1913), pp. 489–500.

† Hazlett, *A theorem on modular covariants*, these Transactions, vol. 21 (1920), pp. 247–254.

‡ Hazlett, *Associated forms in the general theory of modular covariants*, American Journal of Mathematics, vol. 43 (1921), pp. 194–196.

§ *Fundamental systems of protomorphic formal modular seminvariants of binary forms*, read before the Society (Rochester, N. Y.), Sept. 8, 1922. I saw this article in MS. after announcing to the Society (Toronto), on December 28, 1921, the results of Part III of this paper.

This notation, however, has the advantage that it can be used for both algebraic covariants and formal modular covariants; and thus, by its use, the proof of every theorem in this paper is such that it applies both to the modular and the non-modular cases. To help the reader appreciate the diversity of these two interpretations of this theory, the more important theorems are applied to a number of special cases.

As in the symbolic theory of algebraic covariants, it is readily seen that all formal modular covariants are polynomials in a finite number of symbolic expressions having the invariant property. Then are proved several other elementary theorems which assume a simple and familiar form for algebraic covariants, but which, for formal modular covariants, assume a form rather strikingly different.

It is then shown that every isobaric formal modular covariant, C, of the system S is congruent, in the field, to a function of the coefficients of S which is (in a certain general sense) an algebraic covariant of S. When, however, C is not isobaric, it is not congruent to an algebraic covariant of S; but it is, nevertheless, congruent to an algebraic covariant of an enlarged system, S', consisting of the forms of S and other forms obtained from the forms of S by replacing the coefficients of S by their p^nth powers. By various devices, there are expressed in symbolic form the formal modular invariants of a fundamental set for the cubic, modulo 2, and for the quadratic, modulo 3; and the above theorems are verified.

Finally, by using the same theorems, we prove that the set of all formal modular covariants of any binary system, S, with respect to the Galois field, $GF[p^n]$, is such that (1) all syzygies among them are consequences of a finite number of such syzygies; and (2) all formal modular covariants are expressible as polynomials in a finite number of such covariants.

Part II. Symbolic notation

4. **Explanation of the notation.** As in the theory of algebraic covariants, formal modular covariants of a binary form, f, assume a form which is both simple and elegant when the form f is expressed as a product of symbolic linear factors. In the theory of algebraic covariants, it is customary to express a binary form f of degree m as a symbolic mth power; but, for reasons which were explained elsewhere in detail,* it is not possible to express the general form f of degree m as a symbolic mth power and then express all formal modular covariants by means of these symbols for the general Galois field.

For this reason, represent

$$f = a_0 x_1^m + a_1 x_1^{m-1} x_2 + \cdots$$

* *New proofs of certain finiteness theorems in the theory of modular covariants*, these Transactions, vol. 22 (1921), pp. 152–153.

as

$$f = \pi\alpha_x = \pi(\alpha_1 x_1 + \alpha_2 x_2) = (\alpha_1 x_1 + \alpha_2 x_2)(\beta_1 x_1 + \beta_2 x_2) \cdots$$

where it is understood that there are m symbolic factors which are symbolically distinct, and consider its formal modular covariants under the group of transformations (1) where the ξ's and η's are any marks of the Galois field, $GF[p^n]$, of order p^n such that the determinant is not zero in the field. Then, as has been proved elsewhere,* every rational integral covariant (both classic and modular) of f is a polynomial in the symbols $\alpha, \beta, \cdots$ and in the x's. The converse is also true; though, in order that a polynomial in the x's and in the symbols be rational in the a's, (i) it must be symmetric in the m pairs (α_1, α_2), $(\beta_1, \beta_2), \cdots$ and such that each term contains as many α's as β's, as many β's as γ's, etc.; or (ii) it must be a sum of a number of expressions described under (i). Similar remarks apply to covariants of several binary forms. For convenience, any polynomial in the x's and the symbols which has the invariantive property under a transformation of the group will be called a symbolic covariant, even if it do not satisfy the conditions that it be rational in the a's.

5. **Fundamental set of symbolic covariants.** Early in the symbolic theory of algebraic covariants, it is proved that all algebraic invariants are expressible as polynomials in a finite number of symbolic invariants of the type

$$(\alpha\beta) = \alpha_1\beta_2 - \alpha_2\beta_1.$$

Similarly, all algebraic covariants are expressible as polynomials in a finite number of symbolic covariants of the type $(\alpha\beta)$ and $\alpha_x = \alpha_1 x_1 + \alpha_2 x_2$. For formal modular covariants we have a similar theorem, though we have to use new types of symbolic covariants.

For simplicity, we shall first consider covariants of a single form, $f = \pi\alpha_x$. Now it is known that the pair (α_1, α_2) is pseudo-cogredient with $(x_2, -x_1)$.† Accordingly, any formal modular covariant may be regarded, for any purpose not involving the notions of weight or index, as if it were actually an invariant of $m+1$ cogredient pairs, consisting of the pairs of symbols and the pair $(x_2, -x_1)$. But it is known that the formal modular invariants of a number of cogredient points have the finiteness property.‡ Notice that this argument applies equally well to a system of binary forms, and thus we have

THEOREM I. *All rational integral formal modular covariants of a system of binary forms,* $f_i = \pi(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_x\beta_x \cdots$, *with respect to the Galois*

* Preceding reference, top of p. 153.

† That is, the two pairs are cogredient aside from a multiplicative factor which is a power of the determinant of the transformation.

‡ F. B. Wiley, *Proof of the finiteness of the modular covariants of a system of binary forms and cogredient points*, these Transactions, vol. 15 (1914), pp. 431–438.

field $GF[p^n]$, of order p^n, are polynomials in a finite number of symbolic covariants which are polynomials in the x's and in the α's, β's, $\cdots$.

6. **Illustration of Theorem I.** Dickson* has shown that all formal modular invariants of the binary quadratic $f = a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2$, modulo 3, are polynomials in

$$\Delta = a_1^2 - a_0 a_2, \qquad J = a_0 \gamma_0 = a_0 (a_0 + a_1 + a_2)(a_0 + 2a_1 + a_2) a_2,$$
$$B = \beta\gamma_1 \equiv a_1 (a_1^2 - a_0^2)(a_2 - a_0)(a_2^2 - a_1^2),$$
$$\Gamma = (a_0 + a_2)(2a_0 + 2a_1 + a_2)(2a_0 + a_1 + a_2).$$

If we express f in symbolic form (§ 4) as $f = (\alpha_1 x_1 + \alpha_2 x_2)(\beta_1 x_1 + \beta_2 x_2)$, then $a_0 = \alpha_1 \beta_1$, $a_1 \equiv -(\alpha_1 \beta_2 + \alpha_2 \beta_1) \mod 3$, $a_2 = \alpha_2 \beta_2$. Hence it is readily seen that

$$\Delta \equiv (\alpha\beta)^2, \qquad J = (\alpha^3 \alpha)(\beta^3 \beta),$$
$$B \equiv \sum (\alpha^3 \beta)^2 (\beta^3 \beta),$$
$$\Gamma \equiv -N_1 \qquad (\mathrm{mod}\ 3),$$

where

$$N_1 = [(\alpha^3 \beta)^2 (\beta^3 \beta) - (\beta^3 \alpha)^2 (\alpha^2 \alpha)]/(\alpha\beta)^3$$

is one of a fundamental set of formal invariants of the pairs (α_1, α_2) and (β_1, β_2) and is an integral function of the α's and β's. Thus Theorem I is verified for this case. Note, also, that the rational invariants of f are expressible as rational functions in the determinantal symbolic invariants, $(\alpha\beta)$, $(\alpha^3 \beta)$, $(\alpha\beta^3)$, $(\alpha^3 \alpha)$, $(\beta^3 \beta)$.

Part III. Relation between classic and modular covariants

7. **Two kinds of congruences.** In the following sections, we shall frequently have to distinguish between two kinds of congruences, identical congruences and residual congruences. Two polynomials $\phi_1(x)$ and $\phi_2(x)$, with integral coefficients, are said to be *residually congruent* with respect to the field $GF[p^n]$ if $\phi_1(x) \equiv \phi_2(x)$ whenever x is any mark of the field; if, however, $\phi_1(x) \equiv \phi_2(x)$ when x is any number whatsoever, then $\phi_1(x)$ is said to be *identically congruent* to $\phi_2(x)$. For example, $x^{p^n} \equiv x$ is a residual congruence, and $(p+1)x \equiv x$ is an identical congruence.

Moreover, if ϕ_1 and ϕ_2 are polynomials in two sets of variables, the x_i and y_i, then we might have a congruence which is residual as far as the x's are concerned, but which is identical as far as the y's are concerned. For example, when C is a modular covariant, (3) in § 2 is a residual congruence as far as the a's are concerned but is an identical congruence as far as the x's are concerned. If, on the other hand, C is a formal covariant, then (3) is an identical congruence with respect to both the a's and the x's. In each case, however, (3)

* *Madison Colloquium Lectures*, p. 42.

is a residual congruence with respect to the ξ's and η's. Since there is no simple notation which will indicate those variables with respect to which a congruence is residual, we shall indicate them in words directly after each congruence.

8. **Manner in which covariants are transformed.** An algebraic covariant, C, is usually defined as a function such that its transform, C', is given by the formula $C' = D^w C$ where D is the determinant of the general transformation of the group; but it is often defined as a function such that $C' = MC$ where M depends merely on the coefficients of the transformation, and the multiplier is then proved to be a power of D. Now it has been customary thus far to define a formal modular covariant in a manner analogous to the first of these methods, but, in this paper, we shall use the second definition and then prove it is equivalent to the first.

Although the definition of a formal modular covariant says nothing explicitly about the nature of its transform except for values of the ξ's and η's in the field, let us see if this indirectly imposes any restriction on the actual form of the transform before we reduce the exponents of the ξ's and η's by Fermat's theorem. For example, under the transformation (1), $L_1 = \alpha_1^{p^n} \alpha_2 - \alpha_1 \alpha_2^{p^n}$ is replaced by the same function of the primed letters, which is identically equal to

$$D_1 \alpha_1^{p^n+1} + D_2 \alpha_1^{p^n} \alpha_2 - D_3 \alpha_1 \alpha_2^{p^n} + D_4 \alpha_2^{p^n+1},$$

where

$$D_1 = \xi_1^{p^n} \eta_1 - \xi_1 \eta_1^{p^n}, \qquad D_2 = \xi_1^{p^n} \eta_2 - \eta_1^{p^n} \xi_2,$$
$$D_3 = \xi_1 \eta_2^{p^n} - \xi_2^{p^n} \eta_1, \qquad D_4 = \xi_2^{p^n} \eta_2 - \xi_2 \eta_2^{p^n}.$$

It will be observed that each D_i is a formal modular invariant of the two pairs (ξ_1, η_1) and (ξ_2, η_2) when they are transformed cogrediently, and that $D_1 \equiv D_4 \equiv 0$, $D_2 \equiv D_3 \equiv D$, when the ξ's and η's are all in the field.

Let C be any formal covariant of the system S. Then, under (1), C formally —i.e., before any reduction is made modulo p—goes into

$$C(a'; x') = \Sigma D_i(\xi, \eta) P_i(a; x) \tag{4}$$

where each D_i* and each P_i is a polynomial in its arguments. If (1) is followed by the transformation

$$x_i' = e_i x_1'' + f_i x_2'' \qquad (i = 1, 2),$$

where

$$\Delta = e_1 f_2 - e_2 f_1 \not\equiv 0 \quad \text{(in the field)},$$

this is equivalent to applying to the original forms of S the single transformation

$$x_i = \xi_i' x_1'' + \eta_i' x_2'' \qquad (i = 1, 2), \tag{1'}$$

* These multipliers, D_i, will depend on the covariant, C. When C is L, then the D's are as given in the preceding paragraph.

where the two pairs (ξ_1', η_1') and (ξ_2', η_2') are obtained respectively from the pairs (ξ_1, η_1) and (ξ_2, η_2) by applying the transformation of matrix

$$\begin{pmatrix} e_1 & e_2 \\ f_1 & f_2 \end{pmatrix}. \tag{5}$$

Accordingly, if in (4) we replace the a' and the x' by the corresponding a'' and x'', and apply the transformation (5) to the pairs (ξ, η), then we get an equation of the type (4) but formed for the new transformation (1'). Thus, before any reduction is made modulo p,

$$C(a''; x'') = \Sigma D_i(\xi', \eta') P_i(a; x); \tag{4'}$$

and also

$$C(a''; x'') \equiv MC(a'; x') \quad \text{(in the field)}, \tag{6}$$

whenever the e's and f's are in the field. Both (4) and (4') are ordinary equations, but (6) is a congruence in the field, being an identical congruence with respect to the a's and x's. Combining these, we have

$$D_i(\xi', \eta') \equiv MD_i(\xi, \eta)$$

as an identical congruence in the ξ's and η's. Hence we have proved

THEOREM II. *If C is any formal modular covariant of a system, S, of binary forms with respect to the Galois field $GF[p^n]$, of order p^n, then under any linear transformation* (1) *of the group, C is formally replaced by an expression of the form $\Sigma D_i(\xi, \eta) P_i(a; x)$ where each P_i is a polynomial in the a's and x's and each D_i is a formal modular invariant of the two cogredient pairs (ξ_1, η_1) and (ξ_2, η_2).*

In a later section, we shall prove that each D_i which is not congruent to zero when the ξ's and η's are in the field is congruent to a power of D (the determinant of the transformation) whenever the ξ's and η's are marks of the field. For isobaric covariants, see Lemma 4 (§ 10); for pseudo-isobaric covariants, see § 16.

Note that this proof holds equally well for classic covariants, and thus, both for classic covariants and modular covariants, the theory of covariants of a system of binary forms has its foundations in the theory of invariants of two cogredient points. The essential differences between the modular and classic theories are, accordingly, associated with two facts: (1) in the classic case, all invariants of two cogredient points are polynomials in one invariant, the determinant $\xi_1 \eta_2 - \eta_1 \xi_2$; whereas, in the modular case, several other types arise; (2) in the classic case there is no invariant, other than zero itself, which vanishes whenever the ξ's and η's are in the field; whereas, in the modular case, there are invariants of the two pairs which are not identically congruent to zero and yet vanish whenever the ξ's and η's are in the field. From (1)

we see why the transform of a modular covariant is not in general identically a multiple of the original covariant, and from (2) we see why the transform of a modular covariant in general contains terms having no counterpart in the original covariant, before Fermat's theorem is applied. Both remarks are illustrated by L at the beginning of this section.

9. **Remarks on isobarism.** Another fundamental difference between algebraic and modular covariants was mentioned in the introduction and will now be discussed briefly. An algebraic covariant is always isobaric—i.e., all its terms have the same weight.* In fact, an elementary way of determining classic invariants of $f(a; x)$ is to write down a general linear combination of all possible terms having the same degree in the a's and the same weight and then determine what the coefficients of these terms must be in order that the expression be an invariant of f.† But a modular covariant is not in general isobaric, though the weights of any two of its terms can differ at most by a multiple of $p^n - 1$.

These differences are well shown by the algebraic invariants of

$$f = a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2$$

and the formal invariants of the same form, modulo 3. All algebraic invariants of f are polynomials in‡

$$\Delta = a_1^2 - a_0 a_2,$$

and the modular invariants are polynomials in Δ, J, B and Γ (see § 6). It will be noticed that of the modular invariants, only one is isobaric and that is the one which is also an algebraic invariant of f; the weights of any terms of the others differ at most by multiples of $p^n - 1 = 2$.

Isobaric covariants

10. **Preliminary lemmas.**

Lemma I. *Let C be a formal modular covariant of a system, S, of binary forms and a number of cogredient points under a group G with coefficients in the Galois field $GF[p^n]$, of order p^n. Then, if C is isobaric, it is invariant modulo p under any transformation* (1) *where the ξ's and η's are independent variables.*

For, since C is invariant modulo p under the transformation

$$x_1 = x_1' + x_2',$$
$$x_2 = \quad x_2',$$

* In the binary form $f = a_0 x_1^m + a_1 x_1^{m-1} x_2 + \cdots + a_m x_2^m$, the weight of the coefficient a_i is defined to be i, and the variables x_1 and x_2 are assigned the weights 1 and 0 respectively. See any standard book on algebraic invariants, such as Elliott, *Algebra of Quantics*, first edition, pp. 36, 48; Dickson, *Algebraic Invariants*, pp. 31, 38.

† For examples, see Elliott, pp. 125–126; Dickson, pp. 36–37.

‡ Dickson, *Algebraic Invariants*, pp. 48, 84; Elliott, loc. cit., pp. 98–99.

and is isobaric, it is formally invariant, modulo p, under any transformation

$$\begin{aligned} x_1 &= x_1' + kx_2', \\ x_2 &= \quad kx_2', \end{aligned}$$

where k is a general non-zero scalar. Again using the isobarism, it follows from this last that C is invariant modulo p under the transformation

$$\begin{aligned} x_1 &= x_1' + kx_2', \\ x_2 &= \quad x_2'. \end{aligned} \tag{7}$$

Also, C satisfies the condition for covariancy under the transformation

$$\begin{aligned} x_1 &= x_2', \\ x_2 &= x_1', \end{aligned} \tag{8}$$

without using Fermat's theorem. Now, using (7) and (8) in conjunction with the isobarism, we prove the lemma.

LEMMA II. *If a formal modular covariant is isobaric, then every one of its factors is necessarily isobaric.*

For, let C_1 and C_2 be any two formal modular covariants which are not isobaric. Then the totality of terms in $C_1 C_2$ which have the greatest (least) weight is simply the product of the terms of C_1 which have the greatest (least) weight by the terms of C_2 which have the greatest (least) weight. The lemma follows at once.

LEMMA III. *When a formal modular covariant C is isobaric, each of the multipliers $D_i(\xi, \eta)$ of Theorem* II *is also isobaric.*

For, by Theorem II, under any transformation (1), we have (3) holding formally as to the a's, x's, ξ's and η's. Since C is isobaric, it is invariant whenever x_1 or x_2 is multiplied by ρ, where ρ is any non-zero scalar. This means that the right member of (3) is homogeneous in ξ_1 and η_1, and also in ξ_2 and η_2. Being also invariant whenever we interchange the x's, the right member of (3) must be of the same degree in ξ_1 and η_1 that it is in ξ_2 and η_2. Hence, if w_1, w_2, w_3 and w_4 are the degrees of any term of some D_i in ξ_1, η_1, ξ_2 and η_2, respectively, then $w_1 + w_2 = w_3 + w_4$. Similarly, since C is invariant formally whenever x_1' or x_2' is multiplied by ρ, we have $w_1 + w_3 = w_2 + w_4$. Therefore $w_1 = w_4$ and $w_2 = w_3$.

LEMMA IV. *When a formal modular covariant, C, is isobaric, each of the multipliers D_i of Theorem* II *is either identically congruent to zero or to a power of the determinant, D, of the transformation.*

By Lemma III, D_i is a linear combination of terms of the form $\xi_1^k \eta_1^{w-k} \xi_2^{w-k} \eta_2^k$, where w is constant for a given covariant, C; and thus each D_i which is not congruent to zero for every transformation of the group is of the form*

* For if D_i is not congruent to zero when the ξ's and η's are in the field, then it must be equal to unity for the identical transformation and hence must contain a term in $\xi_1^w \eta_2^w$ with coefficient unity.

$$D_i = (\xi_1 \eta_2 - \eta_1 \xi_2)^w + \Delta_i, \tag{9}$$

where Δ_i is divisible by $\eta_1 \xi_2$ and vanishes whenever the ξ's and η's are in the field. But D_i and $\xi_1 \eta_2 - \eta_1 \xi_2$ are both formal modular invariants of the two pairs, (ξ_i, η_i), and are also isobaric; hence Δ_i is likewise. Being divisible by η_1 and ξ_2, Δ_i is divisible by $L_i = \xi_1^{p^n} \eta_i - \xi_i \eta_1^{p^n}$ $(i = 1, 2)$. By Lemma II, this is impossible unless Δ_i is identically zero in the field. Finally, each D_i which is congruent to zero whenever the ξ's and η's are in the field is of the same type as the Δ_i above. Thus the lemma is proved.

11. **Isobaric seminvariants in terms of the roots.** If a formal modular seminvariant I of a system, S, of binary forms with respect to the field $GF[p^n]$ is isobaric, then, by Lemma I, it is invariant, modulo p, under any transformation

$$\begin{aligned} x_1 &= x_1' + M x_2', \\ x_2 &= x_2', \end{aligned} \tag{10}$$

where M is a general scalar. Hence, as in the theory of algebraic seminvariants, we can prove

LEMMA V. *Let I be any isobaric formal modular seminvariant of a system, S, of binary forms with respect to the Galois field, $GF[p^n]$, of order p^n which is of degree d and weight w. Then I is congruent, modulo p, to the product of a_0^d by a symmetric polynomial, S, in the symbolic ratios α_2/α_1, β_2/β_1, $\cdots$ which is homogeneous in these ratios of degree w. Moreover, S is expressible as a polynomial in the differences of these ratios.*

The proof proceeds as in the theory of classic seminvariants, but (for the sake of completeness) we shall reproduce it for the case when S contains only one form. Since I is isobaric, it can not contain a_0 as a factor, by Lemma II. Hence I is a_0^d multiplied by a polynomial, S, of degree d in a_1/a_0, a_2/a_0, $\cdots$, a_m/a_0. Thus S is a polynomial in the elementary symmetric functions of the symbolic ratios

$$A = \frac{\alpha_2}{\alpha_1}, \frac{\beta_2}{\beta_1}, \frac{\gamma_2}{\gamma_1}, \cdots,$$

and is of degree d in each such ratio; and since I is isobaric, S is homogeneous in these symbolic ratios of total degree w.

But, since I is formally unaltered, modulo p, under transformation (10), it is formally unaltered, modulo p, when the symbolic ratios are diminished by any scalar, M. Accordingly, set

$$\frac{\beta_2}{\beta_1} = \left(\frac{\beta_2}{\beta_1} - \frac{\alpha_2}{\alpha_1}\right) + A,$$

and similarly with the other ratios; and thus

$$S = \left(\frac{\alpha_2}{\alpha_1}\right)^k A_k + \left(\frac{\alpha_2}{\alpha_1}\right)^{k-1} A_{k-1} + \cdots + A_0, \tag{11}$$

where each A_i is a polynomial in the differences such as $(\beta_2/\beta_1) - (\alpha_2/\alpha_1)$, $(\gamma_2/\gamma_1) - (\alpha_2/\alpha_1)$, $\cdots$. As in classic invariant theory, this is impossible unless the A_i $(i \geqq 1)$ are all zero in the field. Hence the lemma is proved. The converse is also readily proved, but will not be needed here.

12. **Isobaric invariants in terms of the ratios.** When I is an isobaric formal modular invariant of a binary form, f, then (by Lemma I) its transform under any transformation (1) when the ξ's and η's are general scalars is connected with I by the relation (2), where (2) is an identical congruence as to the a's and thus is essentially of the same type as the corresponding relation for an algebraic invariant of f. Hence, as in the theory of algebraic invariants, we can prove for invariants the result which is a generalization of Lemma V. A similar argument obtains for an invariant of a system of binary forms and hence for covariants of a system of forms.* Thus we have

LEMMA VI. *Let C (a formal modular covariant of a system S of binary forms, f_i) be of degree d_i in the coefficients of f_i and of degree d in the variables, x_1 and x_2, and let it be isobaric of weight w.*

I. *Then C is identically congruent, modulo p, to $\Pi a_0^{d_i} x_2^d$ multiplied by a linear combination, K, of products of the expressions $(x_1/x_2) + (\alpha_2/\alpha_1)$, and of differences of the type $(\beta_2/\beta_1) - (\alpha_2/\alpha_1)$, where the α's and β's may be symbols arising from the same form, f_i, or may be symbols arising from two different forms.*

II. *Moreover, K is homogeneous in the symbols for each form, f_i, and such that each ratio α_2/α_1 for this form, f_i, occurs in exactly d_i factors in each product.*

III. *Finally, K is symmetric, modulo p, in these symbols for each form f_i.*

In applying the results of §§ 11 and 12, the reader must remember that, although the seminvariant (or invariant) is always symmetric in the pairs (α_1, α_2), (β_1, β_2), $\cdots$ in the form in which it first appears, yet the polynomial, K, in the differences $(\beta_2/\beta_1) - (\alpha_2/\alpha_1)$ is not in general symmetric in the ratios $\alpha_2/\alpha_1, \beta_2/\beta_1, \cdots$, but is merely symmetric modulo p. For the binary quadratic $f = a_0 x_1^2 + a_1 x_1 x_2 + a_2 x_2^2 = \alpha_x \beta_x$ has the formal modular invariant $a_1 = \alpha_1 \beta_2 + \alpha_2 \beta_1 \equiv \alpha_1 \beta_1 [(\beta_2/\beta_1) - (\alpha_2/\alpha_1)]$, modulo 2. The first is actually symmetric; but the second is not symmetric, though it is symmetric modulo $p = 2$.

13. **Isobaric formal covariants as algebraic covariants.** From Lemma VI it is evident that, if we express a_0 for each form, f_i, in terms of the symbols, we have

THEOREM III. *Every isobaric formal modular covariant, C, of a system of binary forms is expressible as a polynomial in symbolic covariants of the type $\alpha_x = \alpha_1 x_1 + \alpha_2 x_2$ and in determinantal symbolic invariants of the type*

* For every formal covariant of the system S is a formal modular invariant of the system S' consisting of the forms of S and the additional linear form $l = X_2 x_1 - X_1 x_2$ in which X_1, X_2 have been replaced by x_1, x_2, respectively. The proof of this is the same as for algebraic covariants.

$$(\alpha\beta) = \alpha_1\beta_2 - \alpha_2\beta_1.$$

But it is well known that the product, P, of $x_2^d \Pi a_0^{d_i}$ by any expression, K, in the symbols satisfying the conditions I and II of Lemma VI is an algebraic covariant of the system S of forms, though not necessarily a rational one.* A necessary and sufficient condition that P be rational in the coefficients of f_i is that K be actually symmetric in the ratios β_2/β_1, α_2/α_1, $\cdots$ formed for f_i. This is stronger than condition III of Lemma VI, as is illustrated by the symbolic invariant $(\alpha\beta)$ for the binary quadratic, mod 2 (see § 12). The kth power of $(\alpha\beta)$ is symmetric, mod 2, for any k; but is actually symmetric only when k is even. Thus we have

COROLLARY 1. *Every isobaric formal modular covariant of a system, S, of binary forms with respect to the Galois field, $GF[p^n]$ of order p^n, is congruent, modulo p, to an algebraic covariant of S, either rational or irrational.*

But every irrational algebraic covariant of S is known to be a root of an algebraic equation whose coefficients are rational, algebraic covariants of S, and thus we have

COROLLARY 2. *Every isobaric formal modular covariant of S is congruent, modulo p, to a root of an algebraic equation whose coefficients are rational algebraic covariants of S.*

PSEUDO-ISOBARIC COVARIANTS

14. **Informal discussion.** In the preceding sections, we have considered formal modular covariants that are actually isobaric and have seen that, in many ways, such a covariant is much the same as an ordinary classic covariant. But if a formal modular covariant is not isobaric, what then? For example, consider the formal modular invariants of a point with respect to the field $GF[p^n]$. Any such invariant, I, whose degree, d, is less than $p^n + 1$ must be an ordinary invariant of the point. For, since the weights of any two terms would have to be congruent, mod $p^n - 1$, I would have to be of the form $Ax_1^{p^n-1} + Bx_2^{p^n-1}$ where A and B are constants. But, if the degree is $< p^n$, there is no opportunity to apply Fermat's theorem to the coefficients of the transformation, and thus I must be invariant in the classic sense and therefore identically zero in the field. If $p^n < d \leqq p^{2n}$, it can be shown by a

* Dickson, *Algebraic Invariants*, p. 55, ex. 7; Elliott, loc. cit., p. 93. (Some people do not agree with my statement, and (in justification) say that P in general has no actual meaning in the a's. But any polynomial, P, in the symbols which is not actually symmetric in the symbols, $\alpha_1, \alpha_2, \beta_1, \beta_2, \cdots$ for a particular form, f_i, is a root of an algebraic equation whose coefficients are the elementary symmetric functions of P and the other symbolic invariants obtained from P by interchanging any two pairs of symbols for this form f_i; and hence P is a root of an algebraic equation whose coefficients are polynomials in the coefficients of f_i. Similarly, P is a root of an algebraic equation whose coefficients are polynomials in the coefficients of S, and thus P is an algebraic function of the coefficients of S, though it is not necessarily an explicit algebraic function of these coefficients. Hence why is it not an algebraic covariant of S?)

similar argument that either I is an ordinary invariant of the two cogredient pairs, (x_1, x_2), $(x_1^{p^n}, x_2^{p^n})$, and therefore $I = kL$ where k is a constant and $L = x_1^{p^n} x_2 - x_1 x_2^{p^n}$; or LI is an ordinary invariant of the two cogredient pairs (x_1, x_2) and $(x_1^{p^{2n}}, x_2^{p^{2n}})$, and therefore $LI = k(x_1^{p^{2n}} x_2 - x_1 x_2^{p^{2n}})$. In fact, it is well known that all formal modular invariants of the point (x_1, x_2) are polynomials in L and $Q = (x_1^{p^{2n}} x_2 - x_1 x_2^{p^{2n}})/L$. This might be stated more strikingly by saying that every such invariant is a rational, integral algebraic invariant of three points, or is a quotient of two such invariants. Moreover, a similar statement is true of the formal modular invariants of a pair of cogredient points, (x_1, x_2) and (y_1, y_2).* Is this true more generally?

15. **Preliminary lemmas.** Consider a formal modular invariant, $I = \sum I_j$, of the binary form $f(a; x)$ which is of degree d and which is not isobaric. In particular, let I_0 be a term of I whose weight, w_0, is less than or equal to that of any other term of I, and let I_1 be any other term of I. Then, since I is pseudo-isobaric, the weight of I_1 will be $w_1 = w_0 + l(p^n - 1)$ where l is an integer, positive or zero. When the coefficients of $f(a; x)$ are replaced by the corresponding functions of the symbols $\alpha_1, \alpha_2, \beta_1, \beta_2, \cdots$, I is homogeneous of degree md in these symbols. Also, if I_j is of weight w_j, it is of degree $md - w_j$ in symbols with subscript 1 and of degree w_j in symbols with subscript 2. Their difference, $md - 2w_j$, we shall call the excess, as in classic invariant theory.

LEMMA VII. *The excess of any term of a formal modular covariant is congruent to zero, modulo* $p^n - 1$.†

For, by interchanging the subscripts 1 and 2 throughout I, it follows that I contains a term whose degree in the symbols with subscript 1 is w_j and whose degree in the symbols with subscript 2 is $md - w_j$, and therefore of weight $md - w_j$. Hence $md - w_j \equiv w_j \pmod{p^n - 1}$, and the excess of any term of I is a multiple of $p^n - 1$, say $E_j(p^n - 1)$.

Now I^{p^n} is identically congruent in the field to $\sum I_j^{p^n}$,‡ and thus the latter is a formal modular invariant of $f(a; x)$. It may also be regarded as an invariant of $f(a^{p^n}; x)$, but we shall show that it is better for our purpose to regard it as a simultaneous invariant of $f(a; x)$ and $f(a^{p^n}; x)$.

For in I^{p^n} the term $I_0^{p^n}$ is of excess $E_0 p^n (p^n - 1)$, and hence, if we replace $E_0(p^n - 1)$ factors of the form $\alpha_1^{p^n}$ by as many factors of the form $\bar{\alpha}_1$, the excess of the resulting symbolic expression is $E_0(p^n - 1)$. If $w_0 \geqq E_0$, then by replacing E_0 additional factors of the form $\alpha_1^{p^n}$ by as many factors α_1,

* W. C. Krathwohl, *Modular invariants of two pairs of cogredient variables*, American Journal of Mathematics, vol. 36 (1914), pp. 449–460.

† After finishing this MS., I discovered that Professor Glenn stated and proved this in *Concerning an analogy between formal modular invariants and the class of algebraic invariants called Booleans*, American Journal of Mathematics, vol. 37 (1915), p. 75.

‡ Dickson, *Linear Groups*, p. 15.

the excess of the resulting symbolic expression is zero. $I_0^{p^n}$ is now of degree $E_0 p^n$ in $\bar{\alpha}$'s having subscript 1 and of degree zero in $\bar{\alpha}$'s having subscript 2; also, it is of degree $(w_0 - E_0) p^n$ in α's having subscript 1 and of degree $w_0 p^n$ in the α's having subscript 2. Here and elsewhere in this work, we use the expression "symbols α's" to denote any and all symbols such as $\alpha_1, \alpha_2, \beta_1, \gamma_1, \cdots$; similarly with the expression "symbols $\bar{\alpha}$."

Moreover, if $w_0 \geqq E_0$, the same device will serve to make all terms of I^{p^n} of excess zero and also to make all terms of the same weight, w_0. For the weight of any term, I_ν, of I is $w_0 + k(p^n - 1)$ and hence is of excess $(E_0 - 2k)(p^n - 1)$ which we shall assume positive for convenience. Accordingly, if in $I_\nu^{p^n}$ we replace $(E_0 - 2k)(p^n)$ factors of the form $\alpha_1^{p^n}$ by as many factors $\bar{\alpha}_1$, the result is of excess zero, but of weight $[w_0 + k(p^n - 1)]\, p^n$. Next replace kp^n factors $\alpha_1^{p^n}$ by as many factors $\bar{\alpha}_1$ and replace kp^n factors $\alpha_2^{p^n}$ by as many factors $\bar{\alpha}_2$, and we get an expression whose excess is zero and weight is $w_0 p^n$. This is legal, since $w_0 \geqq E_0 - k + 1$ and $w_0 \geqq k$. In case the excess of I_ν is negative, interchange the subscripts 1 and 2 in the foregoing by an argument similar to that used in Lemma VII. Thus the remark at the beginning of this paragraph is proved.

For convenience, we shall speak of the α's, β's, $\cdots$ (the symbolic coefficients of the linear factors $\alpha_x, \beta_x, \cdots$ of $f(a; x)$) as the *primary symbols*; and we shall speak of the totality of the original symbols $\alpha_1, \alpha_2, \beta_1, \beta_2, \cdots$ together with the new symbols $\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta}_1, \bar{\beta}_2, \cdots$, and any other symbolic coefficients introduced in the above manner as the *secondary symbols.*

We see, accordingly, that every pseudo-isobaric formal modular invariant, I, of $f(a; x)$ is such that a suitable power of I is isobaric in the coefficients of the symbolic linear factors of $f(a; x)$ and of $f(a^{p^n}; x)$, provided $0 \neq w_0 \geqq E_0$. Even if $0 \neq w_0 < E_0$, this argument shows that the excess of I^{p^n} in the secondary symbols is $(E_0 - w_0)(p^n - 1)$ less than that of the original I. But, since the excess is finite, it follows that, by repeating this process at most $[E_0/w_0]$* times, we prove Lemma VIII for the case $w_0 \neq 0$. It will be observed, however, that for this purpose we have to introduce another set of new symbols, say $\bar{\bar{\alpha}}_1, \bar{\bar{\alpha}}_2, \bar{\bar{\beta}}_1, \bar{\bar{\beta}}_2, \cdots$, and possibly still other sets having three or more dashes. These, together with those introduced above, we group under the heading "secondary symbols."

When $w_0 = 0$, we first multiply I by any invariant which is not an absolute invariant and then proceed as above. Thus we have

LEMMA VIII. *Let I be a pseudo-isobaric invariant of a binary form*

$$f(a; x) = \Pi \alpha_x$$

with respect to the Galois field, $GF[p^n]$, of order p^n; and let the lowest weight of all

* Here $[E_0/w_0]$ means the smallest integer which is not less than E_0/w_0.

the terms of I be denoted by w_0. Then, if $w_0 \neq 0$, some finite power of I is congruent to a polynomial in the secondary symbols for f which is isobaric. If $w_0 = 0$, then some power of I is congruent to the quotient of two polynomials in the secondary symbols each of which is isobaric For examples, see Part IV.

16. **Pseudo-isobaric modular covariants as algebraic covariants.** Let J denote that formal modular invariant of f which is obtained as a result of applying the method of Lemma VIII to the pseudo-isobaric invariant I and which is, accordingly, isobaric of weight w in the secondary symbols. If we apply to J the transformation (1), then the transform $J' = J(\alpha')$ is identically equal to

$$J(\alpha'; \bar{\alpha}'; \bar{\alpha}'; \cdots) = \sum D_k(\xi, \eta) P_k(\alpha; \bar{\alpha}; \bar{\alpha}; \cdots) + \tau \tag{12}$$

where each D_k (i) is a formal modular invariant of the two pairs (ξ_1, η_1) and (ξ_2, η_2), (ii) is congruent to zero or to D^w whenever the ξ's and η's are all in the field, (iii) is isobaric of weight w in the ξ's and η's; and where τ is a polynomial in the ξ's and η's and in the secondary symbols which vanishes whenever the dashed symbols are replaced by the proper powers of the corresponding primary symbols. That is, J is not in general an invariant function of the secondary symbols when these symbols are taken as independent, but it is an invariant function when the dashed symbols are replaced by the proper expressions in the primary symbols.

Since J is isobaric in the secondary symbols, Lemmas I–IV apply to J. If we proceed to Lemma V, we see that the proof applies to J with the exception that, since J is not in general rational in the coefficients of $f(a; x)$, $f(a^{p^n}; x)$, $f(a^{p^{2n}}; x)$, $\cdots$, then J is not in general congruent (modulo p) to the product of a_0' by a symmetric polynomial, S, in the secondary symbolic ratios; but, nevertheless, J is congruent (modulo p) to a product of secondary symbols with subscript 1 by a polynomial, S, in the secondary symbolic ratios. Hence, as before, S is expressible as a polynomial in the differences of the secondary symbolic ratios. Thus the proof of Lemma VI applies to J with the restrictions noted.

Hence we prove Theorem III and its corollaries for J, and we have

THEOREM IV. *If I is a formal modular covariant of a system of binary forms with respect to the Galois field $GF[p^n]$ of order p^n, then I^q (where q is a sufficiently large integer) is congruent to one of the two following forms:*

(1) *a polynomial in symbolic covariants of the types α_x, α_X, A_x, $(\alpha\beta)$, $(A\beta)$, etc., where $X = x^{p^n}$ and $A = \alpha^{p^n}$;*

(2) *a quotient of two polynomials described in* (1).

Since these symbolic expressions are (symbolic) algebraic covariants of $f(a; x)$, $f(a^{p^n}; x)$, etc., we have

COROLLARY 1. *Every pseudo-isobaric formal modular covariant, C, of a system S of binary forms with respect to the Galois field $GF[p^n]$ is congruent,*

modulo p, to a symbolic algebraic covariant, C, of the forms of S and certain related forms obtained from the forms of S by replacing each coefficient of these forms by its p^nth power.

In particular, this means that the function J may be chosen in such a way that, under the transformation (1), J goes into J', where J' is equal to (not merely congruent to) $D^w J$, when J is expressed as an isobaric function of the secondary symbols. Note, however, that J is not necessarily rational in the coefficients of the forms of S and the related forms. These remarks are illustrated by examples given in Part IV.

17. **Finiteness of determinantal symbols.** In the preceding section, we proved that all formal modular invariants I of a binary system, S, are congruent, modulo p, to rational functions in symbolic invariants of S of a certain special type called determinantal symbols. That is, I is congruent, modulo p, to a rational function of symbolic invariants of the types $(\alpha\beta)$, $(\alpha^{p^n}\alpha)$ and symbols obtained from these by replacing a pair of symbols by their p^nth powers one or more times. Conceivably, as the degree of I increases, the degree of the requisite determinantal symbols might increase without bound, so that we would need an infinite number of determinantal symbols to obtain all formal modular invariants of S.

But, in Part II, we proved that all formal modular invariants of S are congruent to polynomials in a finite number of symbolic invariants which are formal modular invariants of the (symbolic) coefficients of the symbolic linear factors of the forms of S. Now apply Theorem IV to these symbolic invariants, and we see that all symbolic invariants of S are congruent to rational functions of determinantal symbols. Hence we have

THEOREM V. *All formal modular covariants of a system of binary forms are congruent, in the field, to rational functions of a finite number of symbolic covariants of the types described in Theorem* IV.

In fact, this theorem could have have been proved directly in much the same manner as Theorem IV was proved and then Theorem IV would have followed as a corollary of Theorems I and V.

PART IV. EXPRESSION OF COVARIANTS IN SYMBOLIC FORM

18. **Method for algebraic covariants.** In the symbolic theory of algebraic covariants, after proving that every such covariant is expressible as a polynomial in symbols of the two types, $(\alpha\beta)$ and α_x, there is then given a clear-cut method for expressing any covariant in terms of these symbolic covariants. For completeness, this method will be reproduced here, as we shall need to refer to several of the results.

Let I be any algebraic invariant, of weight w, of a system of forms whose coefficients are the a's; and let $I(a')$ be its transform under (1). Under this

transformation, α_1 and α_2 go into α_ξ and α_η, respectively; and similarly for the β's, γ's, etc. Just as the a's are polynomials in $\alpha_1, \alpha_2, \beta_1, \beta_2, \cdots$, so their transforms are polynomials in $\alpha_\xi, \alpha_\eta, \beta_\xi, \beta_\eta, \cdots$. Hence, in the equation (2),

(i) replace every a' by the corresponding polynomial in α_ξ, α_η, etc.;

(ii) operate on both sides of the resulting equation w times with

$$V = \frac{\partial^2}{\partial \xi_1 \partial \eta_2} - \frac{\partial^2}{\partial \xi_2 \partial \eta_1};$$

(iii) divide both sides of the result of (ii) by $(w+1)!\,w!$.

From the left we get a polynomial in the symbolic invariants of the type $(\alpha\beta)$, and from the right we get $I(a)$.

In proving this, use is made of the following:

LEMMA A.

$$V^w (\xi\eta)^w = (w+1)!\,w!.$$

LEMMA B. *The result of operating by V^w on a product of k terms of type α_ξ and l terms of the type β_η is a sum of terms each containing $k - w$ factors α_ξ, $l - w$ factors β_η and w factors $(\alpha\beta)$.*

19. **Difficulties in modular case.** It is at once evident that this method is not applicable directly to formal modular invariants, since the divisor in (iii) is congruent to zero in the field whenever $w \geqq p - 1$. This difficulty, of course, might be superficial in the sense that it might be due to some accidental peculiarity in the proof and not to any essential characteristic of the problem; but, as will be apparent presently, there is a deep-seated difficulty.

Lemma B is proved by direct verification combined with induction, and thus is lost the full significance of the reason why it is true. Now each symbolic expression α_ξ and each β_η is an invariant of the two points (ξ_1, ξ_2) and (η_1, η_2), which are transformed cogrediently when the original variables are subjected to any linear transformation

$$\begin{aligned} x_1 &= \kappa \bar{x}_1 + \lambda \bar{x}_2, \\ x_2 &= \mu \bar{x}_1 + \nu \bar{x}_2, \end{aligned} \qquad \Delta = \kappa\nu - \lambda\mu \neq 0, \tag{13}$$

thus replacing (1) by

$$\begin{aligned} \bar{x}_1 &= \bar{\xi}_1 x'_1 + \bar{\eta}_1 x'_2, \\ \bar{x}_2 &= \bar{\xi}_2 x'_1 + \bar{\eta}_2 x'_2, \end{aligned} \tag{14}$$

where the $\bar{\xi}$'s and $\bar{\eta}$'s are cogredient with the x's. Moreover, V is an invariantive operator whenever the ξ's and η's are transformed cogrediently. Lemma B is an inevitable consequence of these facts.

If we look at Theorem II (§ 8), we see that in the modular case we have the identity (4) where each D_i is a formal modular invariant of (ξ_1, η_1) and

(ξ_2, η_2) when they are transformed cogrediently; but D_i is not in general a formal modular invariant of the two pairs (ξ_1, ξ_2) and (η_1, η_2), though it is clearly a modular invariant of these pairs. Moreover, the only way to operate on a product of a number of symbols of the type α_ξ and β_η in such a way as (i) to eliminate the ξ's and η's and (ii) to obtain as result a formal modular invariant of the α's, β's, etc., is to use an operator which has the invariantive property when (ξ_1, ξ_2) and (η_1, η_2) are transformed cogrediently. Now the natural analogues of V for the formal modular case are, by Theorem II, those formed on

$$\xi_1^{p^n} \eta_2 - \eta_1^{p^n} \xi_2, \quad \text{etc.},$$

as models—i.e., by replacing each ξ by the operator indicating partial differentiation with respect to that ξ, and similarly with the η's—but they can not give the desired results since these operators are not invariantive when the ξ's are transformed cogrediently with the η's.

20. **Significance of results of Part III.** Such is the essential difficulty of the problem and such is one of the reasons for proving the theorems of Part III. There it is shown that (i) if C is isobaric, then each D_i is an algebraic invariant of the two pairs (ξ_1, η_1) and (ξ_2, η_2) and hence is an invariant of the two pairs (ξ_1, ξ_2) and (η_1, η_2); (ii) if C is not isobaric, then a suitably high power of C, say C^q, is such that, when (2) is formed for C^q, then each D_i is an algebraic invariant of the two pairs (ξ_1, η_1) and (ξ_2, η_2) and hence is an invariant of the pairs (ξ_1, ξ_2) and (η_1, η_2). In view of the remarks of § 19, these results smooth out some of the difficulties.

21. **First method.** Let I be any formal modular invariant of a system S of binary forms with respect to the Galois field, $GF[p^n]$, of order p^n, and let J be a formal modular invariant of S which is obtained from I by applying the method of Lemma VIII and which is isobaric in the secondary symbols. The first method of expressing J in symbolic form is suggested by the proof of Lemma V (§ 11). Although it lacks originality, it has the advantage that it leads at once to a result without exception or qualification. If J is of degree d_i in the coefficients of f_i, divide J by $\pi\alpha_1^{d_i}$ where the α_1 range over all symbols for f_i having subscript 1. Do this for every form f_i. The quotient, S, is a polynomial in the ratios $\alpha_2/\alpha_1, \beta_2/\beta_1, \cdots, \overline{\alpha}_2/\overline{\alpha}_1, \cdots$ with coefficients which are marks of the field. Single out any particular ratio of the primary symbols, say α_2/α_1, and select all terms of S which do not contain α_2/α_1 as a factor. In the resulting expression, replace β_2/β_1 by $(\beta_2/\beta_1) - (\alpha_2/\alpha_1)$, $\alpha_2^{p^n}/\alpha_1^{p^n}$ by $(\alpha_2^{p^n}/\alpha_1^{p^n}) - (\alpha_2/\alpha_1)$, and similarly for all symbolic ratios. The resulting expression, E, is clearly an invariant, since it is a polynomial in the determinantal symbols. Moreover, it is identically congruent, modulo p, to J, for it is simply A_0 in equation (11).

22. **Second method.** It is usually easy to determine by inspection a product, π, of determinantal symbols in such a way that the terms of π which have the greatest excess are among the terms of J which have the greatest excess; and hence, by a finite number of steps, we can express J in symbolic form. No attempt has been made to crystallize this into a definitive, systematic process, though possibly this might be done.

For example, consider $a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2$. Modulo 3, this has the formal modular invariant

$$\begin{aligned} B &= a_1 (a_1 - a_0)(a_1 + a_0)(2a_0 + a_2)(2a_1 + a_2)(a_1 + a_2) \\ &\equiv (\alpha_1\beta_2 + \alpha_2\beta_1)(\alpha_1\beta_2 + \alpha_2\beta_1 + \alpha_1\beta_1)(\alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_1\beta_1) \\ &\qquad \times (-\alpha_1\beta_1 + \alpha_2\beta_2)(\alpha_1\beta_2 + \alpha_2\beta_1 + \alpha_2\beta_2)(\alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_2\beta_2) \\ &\equiv \sum\alpha_1^6\beta_1^3\beta_2^3 - \sum\alpha_1^6\beta_1\beta_2^5 + \sum\alpha_1^4\beta_1^3\alpha_2^2\beta_2^3 \\ &\qquad\qquad - \sum\alpha_1^3\beta_1^2\alpha_2^3\beta_2^4 + \sum\alpha_1^5\alpha_2\beta_2^6 - \sum\alpha_1^3\alpha_2^3\beta_2^6 \\ &\equiv \sum(\alpha^3\beta)^2(\beta^3\beta) \pmod 3. \end{aligned}$$

The terms of greatest excess in B are $\sum\alpha_1^6\beta_1^3\beta_2^3$ which suggests that we use to "kill off" these terms $\sum(\alpha^3\beta)^2(\beta^3\beta)$, since the only way to express $\alpha_1^6\beta_1^3\beta_2^3$ in the secondary symbols $\alpha_1, \alpha_2, \beta_1, \beta_2, \alpha_1^3, \alpha_2^3, \beta_1^3, \beta_2^3$ in such a way that it shall be isobaric in these secondary symbols is $\alpha_1^3\beta_2\,\alpha_1^3\beta_2\,\beta_1^3\beta_2$. Thus we take as π, $(\alpha^3\beta)^2(\beta^3\beta)$. It is then readily discovered that $\sum(\alpha^3\beta)^2(\beta^3\beta)$ accounts for all terms, modulo 3.

23. **Third method.** The preceding method is closely related to the one about to be explained. Although the second method has none of that generality which attracts the artist in pure mathematics, yet it has the advantage in any concrete case, while the present method is to be preferred in any general argument.

Under transformation (1), J is carried into

$$J(a') \equiv D^w J(a) \pmod p \tag{15}$$

where this is an identical congruence in the a's and the coefficients of the transformation. Moreover, it is understood that $J(a')$ is so written that it is isobaric of weight w in the secondary symbols as in Lemma VIII. Thus, if each a' is replaced by its expression in terms of the α_ξ and α_η, $\overline{\alpha}_\xi$ and $\overline{\alpha}_\eta$, etc., the left member of (15) is the sum of a number of terms each of which is a product of w linear functions of the ξ's and w linear functions of the η's.

Now operate on both sides of (15) with V^w as given in (ii), § 18. Since the left member is isobaric in the secondary symbols, we can apply Lemma B. Hence, from the left member we get a sum of a number of terms, each of which is a product of w determinantal symbols such as $(\alpha\beta)$, $(\alpha^{p^n}\beta)$, $(\alpha^{p^n}\alpha)$, etc. From the right we get $(w+1)!\,w!\,J(a)$. Call the result equation E.

If $(w+1)!\,w! \equiv 0$ (modulo p), then the totality of those terms on the right of equation E whose coefficients are not congruent to zero, modulo p, constitute a syzygy, modulo p. Moreover, since the right member of (15) is symmetric, modulo p, in the primary symbols, this syzygy must be symmetric, modulo p, in these symbols and hence congruent to a function which is rational in the coefficients of the ground forms, f. Hence, if we know the necessary syzygies or use discretion, we can replace equation E by one (i) which is equivalent to E and (ii) which has each coefficient congruent to zero, modulo p. If we now cancel p from both sides of this equation, we have an equation, E_1, of the same type as E, but such that the multiplier on the right is not divisible by so high a power of p. Now treat equation E_1 as we treated E above. These steps are all justified, since J is an algebraic invariant of certain forms.

24. **Illustrations.** A fundamental system of formal modular invariants, modulo 3, of the binary quadratic was expressed in symbolic form in § 6.

It is known that a fundamental set of formal modular invariants, modulo 2, of the cubic

$$a_0 x_1^3 + a_1 x_1^2 x_2 + a_2 x_1 x_2^2 + a_3 x_2^3 = \pi\alpha_x = (\alpha_1 x_1 + \alpha_2 x_2)(\beta_1 x_1 + \beta_2 x_2)(\gamma_1 x_1 + \gamma_2 x_2)$$

is

$$K = a_1 + a_2, \qquad \Delta = a_0 a_3 + a_1 a_2, \qquad I = a_0^2 + a_0 K + \delta_{00},$$

$$k = a_0 \delta_{00}, \qquad g = \beta_2^2 + \beta(\Delta + K^2) + (\Delta + \delta_{00})(\beta + a_0 K + K^2)$$

where

$$\beta = a_0 a_1 + a_1^2,$$
$$\delta_{00} = (a_0 + a_1 + a_2 + a_3) a_3.$$

After a little computation, it is readily seen that

$$\Delta \equiv (\alpha\beta)(\beta\gamma)(\gamma\alpha),$$
$$I \equiv \frac{(\alpha^4\alpha)(\beta^4\beta)(\gamma^4\gamma)}{(\alpha^2\alpha)(\beta^2\beta)(\gamma^2\gamma)} + K^2 + \Delta,$$
$$k = (\alpha^2\alpha)(\beta^2\beta)(\gamma^2\gamma).$$

But every attempt to express K as a polynomial in the determinantal symbols is bound to fail, since it is not possible to express K as an isobaric function of the secondary symbols. In fact, K is the invariant which forced the author to realize Theorem IV and the lemmas that lead thereto. It is easy to show that

$$K^2 \equiv \sum(\alpha^2\gamma)(\beta^2\gamma) + (\alpha\beta)(\beta\gamma)(\gamma\alpha).$$

Similarly, it is not possible to express g as a polynomial in determinantal symbols, though (by Theorem IV) a power of g is so expressible.

The author has also expressed some covariants of these two forms in symbolic form, but the amount of numerical work necessary to do this seemed to increase rapidly with the degree of the covariant called J in Part III. In short, the author would not recommend that the theory of Part III be applied in detail to any but the simplest concrete case, unless there should prove to be some powerful systematic procedure not yet discovered.

Part V. Finiteness theorems

25. **Preliminary.** In Part III, we saw that every formal modular covariant of a system, S, of binary forms with respect to the Galois field $GF[p^n]$, of order p^n, is expressible as a polynomial in a finite number of symbolic covariants. Moreover, if the covariant is rational and integral, it must be such that when expressed in terms of these symbolic covariants it is unaltered, modulo p, when any two pairs of primary symbols, (α_1, α_2) and (β_1, β_2), are interchanged. It must be carefully noted, however, that this is not equivalent to saying that it is symmetric in the symbolic covariants.

Using this result, we can prove that the set of all formal modular covariants of S which are rational and integral possesses the finiteness property—i.e., they are all expressible as polynomials in a finite number of these covariants. The proof given below is Hilbert's proof of the finiteness theorem for algebraic covariants, with only one slight modification. The theorem is proved for invariants of a general binary system, S, since the covariants of a system S are coextensive with the invariants of another system S'.

It is readily seen that Hilbert's beautiful proof,* when shorn of non-essentials, depends entirely on the following properties of algebraic invariants:

I. All invariants are polynomials in a finite number of polynomials, P, in a certain finite set of pairs of auxiliary variables. Since these polynomials, P, are invariants under the group, we shall call them elemental invariants. In the case of algebraic invariants, these auxiliary variables are the coefficients of the symbolic linear factors of the forms of S and the elemental invariants are the determinantal symbolic invariants of type $(\alpha\beta)$.

II. If the invariant is rational and integral, it must be such that, when expressed in terms of the elemental invariants, it is unaltered (with respect to the field of definition) when any two pairs of the auxiliary variables are interchanged.

His proof uses the following theorem about diophantine equations: A system of any number of linear diophantine equations has only a finite number of simple sets of solutions, all other solutions being sums of positive multiples of this finite number of sets of solutions.

* *Ueber die Endlichkeit des Invariantensystems für binäre Grundformen*, Mathematische Annalen, vol. 33 (1889), pp. 223–226.

Now Part II proves that these two properties hold for formal modular invariants where the auxiliary variables are the primary symbols $\alpha_1, \alpha_2, \beta_1, \beta_2, \cdots$ and where the elemental invariants are the finite number of symbolic invariants of Theorem I. Moreover, Hilbert's proof makes no use of properties peculiar to a field of characteristic zero except in one place which is readily avoided. Thus his proof applies here.

26. **Finiteness of covariants.** For convenience, let $I_1, I_2, \cdots, I_\mu$ be a fundamental set of symbolic formal modular invariants of S. Then every formal modular invariant of S is expressible in the form

$$I = \sum M_i I_1^{s_1} I_2^{s_2} \cdots I_\mu^{s_\mu} \tag{16}$$

where the exponents, s_j, are positive integers or zero and the coefficients, M_i, are marks of the Galois field, $GF[p^n]$. Necessary and sufficient conditions that I be a rational and integral function of the coefficients of the forms of S are that (i) when the right member of (16) is expanded, it is a function of the pairs $(\alpha_1, \alpha_2), (\beta_1, \beta_2), \cdots$ which is symmetric, modulo p, (ii) each term of this expansion contains the same number of α's as β's, the same number of β's as γ's, etc.

If $d_j(\alpha)$ denotes the degree of I_j in the pair (α_1, α_2), then the last condition gives a set of linear diophantine equations of which two are

$$\begin{aligned} d &= s_1 d_1(\alpha) + s_2 d_2(\alpha) + \cdots + s_\mu d_\mu(\alpha) \\ &= s_1 d_1(\beta) + s_2 d_2(\beta) + \cdots + s_\mu d_\mu(\beta), \end{aligned} \tag{17}$$

there being one such equation for every pair of symbols. By the theory of linear diophantine equations, it follows that every rational integral invariant, I, is expressible in the form

$$I = \sum N_i C_1^{t_1} C_2^{t_2} \cdots C_r^{t_r}, \tag{18}$$

where each C_i is a symbolic invariant of S which has the property (ii) above and where (iii) there is one C_i corresponding to each simple set of solutions of the set of diophantine equations (17), and (iv) the C_i are the same for all invariants, I, of the system S.

Moreover, if σ_i is the number of the symbolic invariants which are not identically congruent, modulo p, and which are obtained from a particular C_i by interchanging any two pairs of symbols, then C_i satisfies an algebraic equation of degree σ_i where the coefficients of the various terms are (aside from sign) the elementary symmetric functions of C_i and its conjugates. Since C_i and therefore each of its conjugates has the property (ii), these coefficients, K_j, are formal modular invariants of S which are rational and integral in the a's. Hence I may be written

$$I = \sum L_j \left(\sum Q_i C_1^{t_1} C_2^{t_2} \cdots C_r^{t_r}\right) \qquad (0 \leqq t_i < \sigma_i) \tag{19}$$

where the Q's are in the field and where the summation in the parentheses is symmetric, modulo p, in the pairs of symbols (α_1, α_2), (β_1, β_2), etc. Moreover, the L_j are polynomials in a finite set of rational, integral formal modular invariants, viz., the K's.

Hence by referring to the sufficient conditions given above that a symbolic invariant be rational and integral, we have

THEOREM VI. *Let S be a system of binary forms with coefficients which are independent variables and let G be the total group of linear transformations on the variables whose coefficients are marks of the Galois field, $GF[p^n]$, of order p^n. Then all formal modular covariants of S under the group G are expressible as polynomials in a finite number of such covariants.*

27. **Finiteness of syzygies.** By using, as in classic invariant theory, Hilbert's useful theorem about an infinite sequence of polynomials, we can show that every syzygy, $S \equiv 0$, among the covariants of the system S is of the form

$$S \equiv \sum H_j S_j,$$

where (i) the H_j are formal modular covariants of S;

(ii) the S_j are polynomials in the formal modular covariants such that $S_j \equiv 0$ is a syzygy among the formal modular covariants of S;

(iii) the S_j are finite in number and are the same for all syzygies.

Thus we have

THEOREM VII. *The syzygies among the formal modular covariants of the system S possess the finiteness property in the sense that all of them are consequences of a finite number of them, $S_1 \equiv 0, S_2 \equiv 0, \cdots, S_k \equiv 0$.*

MOUNT HOLYOKE COLLEGE,
SOUTH HADLEY, MASS.

ON THE MEAN-VALUE THEOREM CORRESPONDING TO A GIVEN LINEAR HOMOGENEOUS DIFFERENTIAL EQUATION*

BY

G. PÓLYA

I consider the linear homogeneous differential expression of order n

$$Lf(x) \equiv f^{(n)}(x) + \phi_1(x) f^{(n-1)}(x) + \phi_2(x) f^{(n-2)}(x) + \cdots + \phi_n(x) f(x).$$

I suppose $\phi_1(x), \phi_2(x), \cdots, \phi_n(x)$ are continuous, $f(x)$ is differentiable n times, the variable x and all the functions in question are real. I then prove some theorems concerning the equation

$$Ly \equiv \frac{d^n y}{dx^n} + \phi_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + \phi_n y = 0 \tag{1}$$

which are, I believe, new, but there is one special case of them which belongs to the most classical part of the differential calculus. This is the case of the simplest equation of the form (1), that is

$$\frac{d^n y}{dx^n} = 0. \tag{2}$$

Denote the Wronskian

$$\begin{vmatrix} f_1(x) & f_1'(x) & \cdots & f_1^{(n-1)}(x) \\ f_2(x) & f_2'(x) & \cdots & f_2^{(n-1)}(x) \\ \cdot & \cdot & \cdot & \cdot \\ f_n(x) & f_n'(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

by $W(f_1(x), f_2(x), \cdots, f_n(x))$. The equation (1) possesses in certain intervals the

Property W. *There exist* $n-1$ *integrals* $h_1(x), h_2(x), \cdots, h_{n-1}(x)$ *of the homogeneous equation* (1) *satisfying the* $n-1$ *inequalities*

$$h_1(x) > 0, \quad W(h_1(x), h_2(x)) > 0, \quad \cdots, W(h_1(x), h_2(x), \cdots, h_{n-1}(x)) > 0 \tag{3}$$

throughout the open interval (a, b).

Determine $n-1$ particular integrals of (1) by such initial conditions that the $n-1$ Wronskians in question are $= 1$ for $x = x_0$; then the inequalities (3) are satisfied in a sufficiently small interval surrounding the point x_0.

* Presented to the Society, October 27, 1923.

Equation (2) possesses the property W in the interval $(-\infty, +\infty)$ since the functions $1, x, x^2, \cdots, x^{n-2}$ are integrals of (2) and we have

$$1 > 0, \quad W(1, x) = 1! > 0, \quad W(1, x, x^2) = 1!\,2! > 0, \quad \cdots.$$

The equation

$$y'' + y = 0 \tag{4}$$

is of the second order. Thus there is only one inequality (3), which is satisfied by the integral $\sin(x - a)$ for $a < x < a + \pi$. No integral of the equation (4) is > 0 throughout an interval (a, b) where $b - a > \pi$. Therefore the property W belongs to (4) in any interval of length $\leqq \pi$ but in no interval of length $> \pi$. An equation $y' + \phi y = 0$ of the first order possesses the property W in any interval where ϕ is continuous.

We assume now (and also in §§ 3–5) *the existence of the property* W. Then we have the following theorems I, II, III:

THEOREM I. *If the function* $f(x)$, *supposed differentiable* n *times, vanishes at* $n + 1$ *points of the interval* (a, b), *then there exists an intermediate point* ξ *such that*

$$Lf(\xi) = f^{(n)}(\xi) + \phi_1(\xi) f^{(n-1)}(\xi) + \cdots + \phi_n(\xi) f(\xi) = 0.$$

Theorem I is a generalization of Rolle's theorem. Apply I to the equation (2) putting $n = 1$; then we have the ordinary theorem of Rolle.

THEOREM II. *There exists one and only one integral* $H(x)$ *of the homogeneous equation* (1) *assuming* n *given values at* n *given points of the interval* (a, b).

Theorem II asserts the possibility of a certain interpolation; in the case of equation (2) and if the n points are distinct, it asserts the possibility of solving the problem which is solved explicitly by Lagrange's formula.

THEOREM III. *Determine an integral* $H(x)$ *of the homogeneous equation* (1) *assuming the same values as* $f(x)$ *at* n *given points of* (a, b); *determine further an integral* $N(x)$ *of the non-homogeneous equation*

$$Ly \equiv \frac{d^n y}{dx^n} + \phi_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + \phi_n y = 1, \tag{5}$$

that vanishes at the n *points in question. There exists a point* ξ *intermediate between these* n *points and an arbitrary point* x *of* (a, b) *such that*

$$f(x) = H(x) + N(x) Lf(\xi).$$

Take the equation (2) and n coincident points; in this particular case Theorem III gives Taylor's formula with Lagrange's remainder-term.

In all the theorems just stated coinciding points are admissible; this will be fully explained in § 1. § 2 enumerates some useful formulas concerning

Wronskians. The proofs of Theorems I, II, III and a few illustrative examples are given in §§ 3–5. In § 6 I try to justify the definition of property W by proving the converse of Theorem II. § 7 deals with a system of several functions of one variable and generalizes Rolle's theorem in another manner.

1. The function $f(x)$ is supposed differentiable n times. We say that $f(x)$ vanishes at k points coinciding with x_0 if

$$f(x_0) = f'(x_0) = \cdots = f^{(k-1)}(x_0) = 0. \tag{6}$$

This statement has a definite meaning if $k \leqq n+1$. If we say that $f(x)$ vanishes at k *and no more* points coinciding with x_0, then we add to (6) the condition

$$f^{(k)}(x_0) \neq 0,$$

and we suppose $k \leqq n$. If $f(x)$ vanishes at

$$m_1, m_2, m_3, \cdots, m_l \tag{7}$$

points coinciding with

$$x_1, x_2, x_3, \cdots, x_l \tag{8}$$

respectively where

$$a < x_1 < x_2 < x_3 < \cdots < x_l < b \tag{9}$$

then we say that $f(x)$ vanishes at $m_1 + m_2 + \cdots + m_l$ points of the interval (a, b). If we say that $f(x)$ and $\phi(x)$ assume the same values at n points of (a, b) we mean that $f(x) - \phi(x)$ vanishes at n points of (a, b). Theorem II may be stated more explicitly thus: if there are given l points (8) satisfying (9), l positive integers (7) such that $m_1 + m_2 + \cdots + m_l = n$, and finally l systems of values

$$\begin{array}{l} y_1, y_1', y_1'', \cdots, y_1^{(m_1-1)}, \\ y_2, y_2', y_2'', \cdots, y_2^{(m_2-1)}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_l, y_l', y_l'', \cdots, y_l^{(m_l-1)}, \end{array}$$

then there exists an integral $H(x)$ of (1) satisfying the conditions

$$\begin{array}{llll} H(x_1) = y_1, & H'(x_1) = y_1', & \cdots, & H^{(m_1-1)}(x_1) = y_1^{(m_1-1)}, \\ H(x_2) = y_2, & H'(x_2) = y_2', & \cdots, & H^{(m_2-1)}(x_2) = y_2^{(m_2-1)}, \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ H(x_l) = y_l, & H'(x_l) = y_l', & \cdots, & H^{(m_l-1)}(x_l) = y_l^{(m_l-1)}, \end{array} \tag{10}$$

and $H(x)$ is completely determined by these conditions.

A point ξ satisfying the inequality $x_1 < \xi < x_l$ is said to be *intermediate* between the points (8); so far we have supposed $l > 1$. If $l = 1$, that is, if all the points in question coincide with one point x_1, then $\xi = x_1$ is the only *intermediate* point.

I enumerate a few evident propositions used in what follows. The functions $u, u_1, \cdots, u_k$ are supposed continuous and differentiable as many times as required to calculate the expression

$$(11)\qquad \begin{aligned} &u_k \frac{d}{dx}\left(u_{k-1}\frac{d}{dx}\left(u_{k-2}\cdots\frac{d}{dx}\left(u_1\frac{d}{dx}(uf)\right)\cdots\right)\right) \\ &= u_k\frac{d}{dx}u_{k-1}\frac{d}{dx}u_{k-2}\cdots\frac{d}{dx}u_1\frac{d}{dx}uf = uu_1u_2\cdots u_k\frac{d^kf}{dx^k}+\cdots. \end{aligned}$$

1. *If $f(x)$ vanishes at k (and no more) points coinciding with x_0, then $f'(x)$ vanishes at $k-1$ (and no more) points coinciding with x_0.*

2. *If $f(x)$ vanishes at k points coinciding with x_0, then $u(x)f(x)$ vanishes at k (or more) points coinciding with x_0.*

3. *Suppose $f(x)$ vanishes at k and no more points coinciding with x_0, and suppose further that $u, u_1, u_2, \cdots, u_k$ are different from zero for $x = x_0$. Then the expression* (11) *is $\neq 0$ for $x = x_0$.*

4. *Suppose $f(x)$ vanishes at $k+1$ points. Then, between these $k+1$ points, there exists an intermediate point ξ at which the expression* (11) *vanishes.*

We obtain 4 by repeated application of Rolle's theorem and of Theorems 1, 2.

2. Let $c_1, c_2, \cdots, c_{n-1}$ be constants; then we have

$$(12)\qquad \begin{aligned} &W(f_1, f_2, \cdots, f_{n-1}, c_1f_1 + c_2f_2 + \cdots + c_{n-1}f_{n-1} + f_n) \\ &\qquad = W(f_1, f_2, \cdots, f_{n-1}, f_n). \end{aligned}$$

Suppose the function u is differentiable $n-1$ times; then we have

$$W(uf_1, uf_2, \cdots, uf_n) = u^n W(f_1, f_2, \cdots, f_n).$$

In particular put $u = 1/f_1$; we obtain

$$(13)\qquad W\left[\left(\frac{f_2}{f_1}\right)', \left(\frac{f_3}{f_1}\right)', \cdots, \left(\frac{f_n}{f_1}\right)'\right] = \frac{W(f_1, f_2, \cdots, f_n)}{f_1^n}.$$

For the derivative of a Wronskian we have

$$(14)\qquad W'(f_1, f_2, \cdots, f_n) = |f_k, f_k', \cdots, f_k^{(n-2)}, f_k^{(n)}|_{k=1, 2, \cdots, n}.$$

Using (14) we obtain by the usual formula for a minor of the adjoint determinant

$$\begin{aligned} &W(f_1, \cdots, f_{n-2}, f_{n-1})\,W'(f_1, \cdots, f_{n-2}, f_n) \\ &\qquad - W'(f_1, \cdots, f_{n-2}, f_{n-1})\,W(f_1, \cdots, f_{n-2}, f_n) \\ &\qquad\qquad = W(f_1, \cdots, f_{n-2})\,W(f_1, \cdots, f_{n-2}, f_{n-1}, f_n), \end{aligned}$$

$$(15)\qquad \begin{aligned} &\frac{d}{dx}\frac{W(f_1, \cdots, f_{n-2}, f_n)}{W(f_1, \cdots, f_{n-2}, f_{n-1})} \\ &\qquad = \frac{W(f_1, \cdots, f_{n-2})\,W(f_1, \cdots, f_{n-2}, f_{n-1}, f_n)}{(W(f_1, \cdots, f_{n-2}, f_{n-1}))^2}. \end{aligned}$$

If $h_1(x), h_2(x), \cdots, h_n(x)$ are n linearly independent integrals of (1), we have

$$Ly \equiv \frac{W(h_1, h_2, \cdots, h_n, y)}{W(h_1, h_2, \cdots, h_n)}, \tag{16}$$

$$W(h_1, h_2, \cdots, h_n) = ce^{-\int \phi_1(x)dx}, \tag{17}$$

where c is a constant. Put

$$\begin{aligned} y &= Y_0, & W(h_1, y) &= Y_1, \\ W(h_1, h_2, y) &= Y_2, & \cdots, \quad W(h_1, h_2, \cdots, h_n, y) &= Y_n, \\ 1 &= W_0, & h_1 &= W_1, \\ W(h_1, h_2) &= W_2, & \cdots, \quad W(h_1, h_2, \cdots, h_n) &= W_n. \end{aligned}$$

Applying (15) we obtain

$$\frac{d}{dx}\frac{Y_0}{W_1} = \frac{W_0 Y_1}{W_1^2}, \quad \frac{d}{dx}\frac{Y_1}{W_2} = \frac{W_1 Y_2}{W_2^2}, \quad \cdots, \quad \frac{d}{dx}\frac{Y_{n-1}}{W_n} = \frac{W_{n-1} Y_n}{W_n^2}.$$

Combining these formulas with (16) we find finally

$$Ly \equiv \frac{W_n}{W_{n-1}} \frac{d}{dx} \frac{W_{n-1}^2}{W_{n-2} W_n} \cdots \frac{d}{dx} \frac{W_2^2}{W_1 W_3} \frac{d}{dx} \frac{W_1^2}{W_0 W_2} \frac{d}{dx} \frac{y}{W_1}. \tag{18}$$

(18) is the usual decomposition of a linear homogeneous differential expression into a "product" of differential expressions of the first order. The proof here given is perhaps more direct than the usual.*

3. The property W implies the non-vanishing of $W_1, W_2, \cdots, W_{n-1}$ appearing in formula (18). W_n vanishes nowhere, by virtue of (17). Applying Theorem 4 of § 1 to the decomposition (18) we get at once Theorem I.

Apply now Theorem I to $f(x) - H(x)$ instead of $f(x)$, where $H(x)$ is an integral of the homogeneous equation (1). Then $LH(x) = 0$ and we obtain the corollary

THEOREM I*. *If the function $f(x)$, supposed differentiable n times, assumes at $n+1$ points of the interval (a, b) the same values as an integral of the homogeneous equation* (1), *there exists an intermediate point ξ such that*

$$Lf(\xi) = f^{(n)}(\xi) + \phi_1(\xi) f^{(n-1)}(\xi) + \cdots + \phi_n(\xi) f(\xi) = 0.$$

Theorem I* was incidentally remarked by H. Poincaré† in the case of an equation of the second order where the property W is particularly simple, including only one inequality. Poincaré's proof is very different from the proof given here and more complicated. I obtained the results presented in

* Cf. L. Schlesinger, *Handbuch der Theorie der linearen Differentialgleichungen*, vol. I, p. 52, formula (14).

† L'Intermédiaire des Mathématiciens, vol. 1 (1894), pp. 141–144 and also pp. 69, 127, 172, 216.

this paper in clearing up and working out the original remark made by Poincaré.

Another corollary to Theorem I is the following

THEOREM I**. *No integral $N(x)$ of the non-homogeneous equation* (5) *vanishes at more than n points of the interval (a, b).*

Suppose, if possible, that $N(x)$ vanishes at $n+1$ points; then we have by Theorem I

$$LN(\xi) = 0$$

at an intermediate point ξ; now, actually, $LN(\xi) = 1$ and thus Theorem I** is proved.

4. The general integral of (1) has the form

$$H(x) = c_1 h_1(x) + c_2 h_2(x) + \cdots + c_n h_n(x),$$

where $c_1, c_2, \cdots, c_n$ are constants. The conditions (10) form a system of n linear non-homogeneous equations with the aid of which we have to determine $c_1, c_2, \cdots, c_n$. We have to prove that this system is neither incompatible nor indeterminate. Thus we have to prove that the only solution of the *homogeneous* system obtained from (10) by putting

$$y_1 = y_1' = y_2'' = \cdots = y_i^{(m_i-1)} = 0 \tag{19}$$

is the identically vanishing solution.

Suppose, if possible, that there is a solution of the homogeneous system (10) (19) which does not vanish identically; denote it by

$$H(x) = \gamma_1 h_1(x) + \gamma_2 h_2(x) + \cdots + \gamma_k h_k(x) + \cdots + \gamma_n h_n(x)$$

where

$$\gamma_k \neq 0, \qquad \gamma_{k+1} = \gamma_{k+2} = \cdots = \gamma_n = 0.$$

Then $h_k(x)$ assumes at $n \geqq (k-1)+1$ points the same values as an integral of the equation

$$\frac{W(h_1, h_2, \cdots, h_{k-1}, y)}{W(h_1, h_2, \cdots, h_{k-1})} = 0$$

of order $k-1$. Then there exists, by virtue of I*, an intermediate point ξ at which

$$W(h_1(\xi), h_2(\xi), \cdots, h_{k-1}(\xi), h_k(\xi)) = 0.$$

This last equation is excluded if we assume the existence of the property W and thus Theorem II is proved.

Example. Assume that the algebraic equation

$$x^n + c_1 x^{n-1} + \cdots + c_n = 0$$

has n different real roots $a_1, a_2, \cdots, a_n$, where

$$a_1 < a_2 < a_3 < \cdots < a_n. \tag{20}$$

Then the linear homogeneous equation with constant coefficients

$$y^{(n)} + c_1 y^{(n-1)} + c_2 y^{(n-2)} + \cdots + c_n y = 0$$

has the integrals

$$e^{a_1 x},\ e^{a_2 x},\ \cdots,\ e^{a_n x}, \tag{21}$$

and possesses the property W in the interval $(-\infty, +\infty)$ since the Wronskians

$$e^{a_1 x},\ W(e^{a_1 x}, e^{a_2 x}),\ \cdots,\ W(e^{a_1 x}, e^{a_2 x}, \cdots, e^{a_n x})$$

divided by

$$e^{a_1 x},\ e^{(a_1+a_2)x},\ \cdots,\ e^{(a_1+a_2+\cdots+a_n)x}$$

respectively reduce to positive constants which may be calculated by Vandermonde's formula. Therefore we can interpolate by a suitably chosen linear combination of the n integrals (21) n arbitrary values at the n points $x_1, x_2, \cdots, x_n$ where

$$x_1 < x_2 < x_3 < \cdots < x_n. \tag{22}$$

This fact is equivalent to the statement that

$$\begin{vmatrix} e^{a_1 x_1} & e^{a_1 x_2} & \cdots & e^{a_1 x_n} \\ e^{a_2 x_1} & e^{a_2 x_2} & \cdots & e^{a_2 x_n} \\ \cdot & \cdot & \cdot & \cdot \\ e^{a_n x_1} & e^{a_n x_2} & \cdots & e^{a_n x_n} \end{vmatrix} \neq 0. \tag{23}$$

Thus we have proved (23) as a particular case of Theorem II. It may be shown directly in a slightly different manner that the determinant (23) is > 0 if (20) (22) are satisfied. I add that the property W belongs to any equation whatever with constant coefficients in the interval $(-\infty, +\infty)$ if the roots of the corresponding algebraic equation are all real, *whether distinct or not*. We get a simple proof of this from (17).

Any integral $N(x)$ of the non-homogeneous equation (5) is of the form

$$N(x) = N_0(x) + H(x).$$

Here $N_0(x)$ is a particular integral of the non-homogeneous equation (5) and $H(x)$ is the general integral of the homogeneous equation (1); $H(x)$ contains n arbitrary constants which we can adapt to any conditions of the form (10), by virtue of Theorem II. Thus we obtain the corollary

THEOREM II*. *There exists one and only one integral $N(x)$ of the non-homogeneous equation* (5) *assuming n given values at n given points of the interval* (a, b).

5. I now prove Theorem III. The possibility of determining $H(x)$ according to the enunciated conditions follows from Theorem II and the possibility

of determining $N(x)$ follows from II$*$. Let x_0 be a point of (a, b) but not one of the n points mentioned at which $f(x) - H(x)$ and $N(x)$ vanish. Then we have, by Theorem I$**$, $N(x_0) \neq 0$. Therefore it is possible to find a constant C satisfying the equation

$$f(x_0) = H(x_0) + CN(x_0). \tag{24}$$

Thus the function

$$f(x) - H(x) - CN(x)$$

vanishes at $n + 1$ *points* of (a, b). There exists, by Theorem I, an intermediate point ξ at which

$$L(f(\xi) - H(\xi) - CN(\xi)) = Lf(\xi) - 0 - C \cdot 1 = 0,$$

that is,

$$C = Lf(\xi).$$

Substituting this value in (24) we have Theorem III.*

I will now give some concrete cases of Theorem III with immediate applications.

Example 1. Let x_1 be the point at which we interpolate the differentiable function $f(x)$ by an integral of the equation

$$y' + y = 0.$$

Adopting the notation of Theorem III we have

$$H(x) = f(x_1)e^{x_1 - x}, \qquad N(x) = 1 - e^{x_1 - x},$$

and consequently for $x > x_1$

$$f(x) = f(x_1)e^{x_1 - x} + (1 - e^{x_1 - x})(f(\xi) + f'(\xi)) \tag{25}$$

where $x_1 < \xi < x$. From (25) we get at once the proposition †

Suppose $f(x)$ *differentiable for* $x > 0$; *then*

$$\lim_{x \to +\infty} (f(x) + f'(x)) = c$$

implies $\lim f(x) = c$.

We may assume $c = 0$ (and apply the proof to $f(x) - c$ instead of to $f(x)$). Choose first x_1 sufficiently great so that $|f(x) + f'(x)| < \epsilon/2$ for $x > x_1$; then choose x_2, $x_2 > x_1$, such that $|f(x_1)|e^{x_1 - x_2} < \epsilon/2$. Finally we get by (25) for $x > x_2$

$$|f(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

* The particular case corresponding to (2) was given by Ch. Hermite for points coinciding in an arbitrary manner. See *Œuvres*, vol. III, pp. 432–443, and T. J. Stieltjes, *Œuvres*, vol. I, pp. 47–60. For a more general point of view see G. D. Birkhoff, these Transactions, vol. 7 (1906), pp. 107–136.

† G. H. Hardy, Quarterly Journal of Mathematics, vol. 35 (1903), pp. 31–33. The proof given by O. Perron, Mathematische Zeitschrift, vol. 6 (1920), p. 159, supposes *more* than the derivability of $f(x)$.

By the same argument we may prove an analogous result for any equation with constant coefficients if the roots of the corresponding algebraic equation are real and negative.*

Example 2. In dealing with the equation

$$y'' + y = 0 \tag{4}$$

we may choose for (a, b) any interval of length π. The two points at which we interpolate may be different or not. Interpolating at two distinct points x_1, x_2, where

$$x_1 < x_2 < x_1 + \pi, \tag{26}$$

we find the formula

$$f(x) = \frac{f(x_1)\sin(x_2 - x) + f(x_2)\sin(x - x_1)}{\sin(x_2 - x_1)} + \left(1 - \frac{\cos\left(x - \frac{x_1 + x_2}{2}\right)}{\cos\frac{x_2 - x_1}{2}}\right)(f''(\xi) + f(\xi)) \tag{27}$$

valid for $x_2 - \pi < x < x_1 + \pi$; ξ is an interior point of the least interval containing x_1, x_2, x. Interpolating at two points coinciding with x_1 we find

$$f(x) = f(x_1)\cos(x - x_1) + f'(x_1)\sin(x - x_1) + (1 - \cos(x - x_1))(f''(\xi) + f(\xi)) \tag{28}$$

valid for $x_1 - \pi < x < x_1 + \pi$; ξ lies between x_1 and x. We see that restriction (26) is essential for the validity of (27), $x_2 = x_1 + \pi$ being generally inadmissible; this point will be cleared up by Theorem IV.

In order to give an application suppose $f(x)$ is an integral of the equation

$$y'' + \phi(x)y = 0, \tag{29}$$

where

$$\phi(x) > 1; \tag{30}$$

suppose further that $f(x_1) > 0$. We have by (29)

$$f''(\xi) + f(\xi) = -(\phi(\xi) - 1)f(\xi);$$

thus we may put (28) in the form

$$f(x) = A\sin(x - a) - (1 - \cos(x - x_1))(\phi(\xi) - 1)f(\xi) \tag{31}$$

* O. Perron, Mathematische Zeitschrift, vol. 6 (1920), pp. 161–163.

where

$$A \sin (x_1 - a) = f(x_1) > 0, \qquad A \cos (x_1 - a) = f'(x_1). \tag{32}$$

The situation shown by (30) (31) (32) is this: the curve $A \sin (x - a)$ meets the curve $f(x)$ at the point x_1, see (32); the curve $A \sin (x - a)$ runs on the upper side of the curve $f(x)$ *as far* as $f(x)$ runs on the upper side of the x-axis, see (31) (30); but the curve $A \sin (x - a)$ cuts the x-axis at two points between which x_1 lies and whose distance apart is π; hence $f(x)$ reaches the x-axis necessarily *before* $A \sin (x - a)$ *cuts it* and we have the theorem

The distance between two consecutive zeros of any integral of the equation (29) *satisfying the condition* (30) *is* $< \pi$.

This is a particular case of a well known classical theorem given by Sturm* that can be proved generally by the argument I have used in the foregoing concrete case.

Any integral $f(x)$ of the equation (29) has the following property: the inequality

$$f(x) > 0 \tag{33}$$

implies

$$f''(x) + f(x) < 0. \tag{34}$$

Consider the curve of which the equation is

$$r = \frac{1}{f(\vartheta)} \tag{35}$$

in a system of polar coördinates r, ϑ with origin O. The inequality (34) expresses that the curve (35) appears from O to be *convex*, as we may see, e.g., by (27). The result just proved may be formulated geometrically as follows:

A curve appearing from a given point O to be everywhere convex subtends an angle $< \pi$ *at* O.

H. Poincaré was led to the remark above mentioned (§ 3) by constructing an analytical demonstration of this geometrically evident fact.

6. The converse of Theorem II is also true: if the interpolation of any n values by an integral of (1) is possible at any n points of an interval, then the equation (1) possesses the property W in that interval. The possibility of interpolation, that is, the *existence* of a solution of the non-homogeneous system (10) with arbitrary $y_1, y_1', \cdots, y_1^{(m_1-1)}$, is equivalent to the *non-existence* of a non-vanishing solution of the corresponding homogeneous system; we used this equivalence in § 4. Thus the most simple form of the converse to Theorem II is the following:

* Cf. Maxime Bôcher, *Méthodes de Sturm*, Paris, 1917, p. 52.

THEOREM IV. *If, except the identically vanishing integral, there exists no integral of equation* (1) *vanishing at n points of the interval* $a \leqq x < b$, *then there exist* $n-1$ *integrals* $h_1(x), h_2(x), \cdots, h_{n-1}(x)$ *such that the Wronskians*

(36) $\quad h_1(x),\ W(h_1(x), h_2(x)),\ \cdots,\ W(h_1(x), h_2(x), \cdots, h_{n-1}(x))$

vanish nowhere in the interval $a < x < b$.

The signs of the Wronskians are of no importance provided that they are constant. If the Wronskians (36) do not vanish we can make them positive by substituting $-h_\nu(x)$ for $h_\nu(x)$ for certain values of ν.

Determine n integrals $h_1(x), h_2(x), \cdots, h_n(x)$ satisfying the initial conditions

$$\begin{aligned}
&h_1(a) = h_1'(a) = h_1''(a) = \cdots = h_1^{(n-3)}(a) = h_1^{(n-2)}(a) = 0, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad h_1^{(n-1)}(a) = 1, \\
&h_2(a) = h_2'(a) = h_2''(a) = \cdots = h_2^{(n-3)}(a) = 0,\ h_2^{(n-2)}(a) = 1, \\
&h_{n-2}(a) = h_{n-2}'(a) = 0,\quad h_{n-2}''(a) = 1, \\
&h_{n-1}(a) = 0,\quad h_{n-1}'(a) = 1, \\
&h_n(a) = 1.
\end{aligned} \tag{37}$$

These integrals are linearly independent because their Wronskian $= \pm 1$ for $x = a$; $h_1(x)$ vanishes at $n-1$ points coinciding with a; by hypothesis there exists no nth point in the interval $a \leqq x < b$ at which $h_1(x)$ vanishes; therefore $h_1(x) \neq 0$ for $a < x < b$. Adopt the notation

$$W(h_1, h_2, \cdots, h_k) = W_k$$

that we used in § 2 and suppose that we have already proved

$$W_1 \neq 0,\quad W_2 \neq 0,\quad \cdots,\quad W_k \neq 0 \tag{38}$$

for $a < x < b$, $k < n$.

Let $c_1, c_2, \cdots, c_k$ be constants and put

$$c_1 h_1(x) + c_2 h_2(x) + \cdots + c_k h_k(x) + h_{k+1}(x) = h(x).$$

$h(x)$ does not vanish identically, and therefore it vanishes at no more than $n-1$ points contained in the interval $a \leqq x < b$. Now $h(x)$ vanishes at $n-k-1$ points coinciding with a by virtue of the conditions (37). Determine the constants $c_1, c_2, \cdots, c_k$ in such a manner that $h(x)$ vanishes at k points coinciding with x_0 where $a < x_0 < b$; this is possible because the determinant of the k equations in question

$$W(h_1(x_0), h_2(x_0), \cdots, h_k(x_0)) \neq 0,$$

by the hypothesis of complete induction, that is, by (38). Beyond the

$$(n-k-1) + k = n-1$$

zeros already mentioned, $h(x)$ has no other zeros for $a \leqq x < b$. In particular $h(x)$ *vanishes at no more than* k *points coinciding with* x_0. Consequently, by virtue of hypothesis (38) and of the Theorem 3 of § 1,

$$(39)\qquad \frac{W_k^2}{W_{k-1}}\frac{d}{dx}\frac{W_{k-1}^2}{W_{k-2}W_k}\cdots\frac{d}{dx}\frac{W_2^2}{W_1W_3}\frac{d}{dx}\frac{W_1^2}{W_0W_2}\frac{d}{dx}\frac{h}{W_1} = W(h_1, h_2, \cdots, h_k, h) = W(h_1, h_2, \cdots, h_k, h_{k+1}) = W_{k+1}$$

does not vanish for $x = x_0$; we obtain (39) from (15) (12) in the same manner as we obtained (18). The point x_0 was arbitrarily chosen in the interval $a < x < b$ and thus Theorem IV is proved.

7. Another demonstration of the Theorems I, II, III essentially different from that given in §§ 1–5 is based on the following theorem,

THEOREM V. *Suppose all the functions* $h_1(x), h_2(x), \cdots, h_n(x), f(x)$ *possess derivatives of order* n *throughout the interval* $a < x < b$ *and satisfy the following* n *inequalities*:

$$h_1(x) > 0, \qquad W(h_1(x), h_2(x)) > 0, \qquad \cdots, \qquad W(h_1(x), h_2(x), \cdots, h_n(x)) > 0.$$

Suppose also that the l *points* $x_1, x_2, \cdots, x_l$ *and the corresponding positive integers* $m_1, m_2, \cdots, m_l$ *satisfy the conditions*

$$a < x_1 < x_2 < \cdots < x_l < b, \qquad m_1 + m_2 + \cdots + m_l = n + 1.$$

Consider the determinant of order $n + 1$

$$(40)\qquad \left|h_k(x_1), h_k'(x_1), \cdots, h_k^{(m_1-1)}(x_1), h_k(x_2), h_k'(x_2), \cdots, h_k^{(m_2-1)}(x_2), \cdots, h_k(x_l), \cdots, h_k^{(m_l-1)}(x_l)\right|,$$

the n *first rows being obtained by putting successively* $k = 1, 2, 3, \cdots, n$ *and the last row by substituting* $f(x)$ *for* $h_k(x)$.

There exists an intermediate point ξ, $x_1 < \xi < x_l$, *such that the value assumed by the Wronskian at the point* ξ, *that is,*

$$W(h_1(\xi), h_2(\xi), \cdots, h_n(\xi), f(\xi)),$$

is $>$, $=$ *or* < 0, *according as the determinant* (40) *is* $>$, $=$ *or* < 0.

Theorem V may be compared to a well known determinantal theorem of H. A. Schwarz.* Theorem V may be proved by complete induction; its validity for the $(n-1)+1$ functions

* H. A. Schwarz, *Abhandlungen*, vol. II, pp. 296–302; T. J. Stieltjes, *Œuvres*, vol. II, pp. 110–123. The determinant (40) has been considered by H. M. Morse, G. A. Pfeiffer, and G. M. Green from a different point of view; see Bulletin of the American Mathematical Society, vol. 23 (1916), pp. 114–122.

$$\left(\frac{h_2(x)}{h_1(x)}\right)', \left(\frac{h_3(x)}{h_1(x)}\right)', \cdots, \left(\frac{h_n(x)}{h_1(x)}\right)', \left(\frac{f(x)}{h_1(x)}\right)'$$

implies its validity for the $n+1$ functions $h_1(x), h_2(x), \cdots, h_n(x), f(x)$, as may be shown by (13) and by Rolle's theorem. I had originally based my demonstration of Theorems I, II, III on Theorem V. I was led to the treatment of the subject I finally adopted by a kind remark made by Professor H. Weyl.

EIDG. TECHNISCHE HOCHSCHULE,
ZURICH, SWITZERLAND.

ERRATA, VOLUME 24

J. F. RITT, *On algebraic functions which can be expressed in terms of radicals.* Page 21, lines 30 and 33, for "n^2" read "n".